Ph.D. Dissertation

ECONOMIC DETERMINANTS OF INDIA`S INVESTMENT CLIMATE

BY

MANUEL DE ORTE RAMÍREZ
BA in economics

Universidad Nacional de Educación a Distancia. Department of Economic Analysis II.
Madrid, Spain. 2010.

Ph.D. Dissertation

ECONOMIC DETERMINANTS OF INDIA`S INVESTMENT CLIMATE

BY
Manuel de Orte Ramírez
BA in economics, Universidad Complutense de Madrid.
MA in industrial organization and markets, Universidad Carlos III de Madrid.

Universidad Nacional de Educación a Distancia. Department of Economic Analysis II.
Madrid, Spain. 2010.

Universidad Nacional de Educación a Distancia. Department of Economic Analysis II. Madrid, Spain. 2010.

PhD Dissertation prepared by: Manuel de Orte Ramírez, BA in economics, Universidad Complutense de Madrid.

Supervisor: Álvaro Escribano Sáez, Department of Economics, Universidad Carlos II de Madrid.

Tutor of the dissertation: Ana Martín Marcos, Department of Economic Analysis II, Universidad Nacional de Educación a Distancia.

Dedication

To my parents, Inmaculada and Manuel, and my brother, Javier.

Acknowledgments

My deepest gratitude is to my advisor. This research project would not have been possible without my supervisor, Prof. Dr. Álvaro Escribano who was abundantly helpful and offered invaluable assistance, support and guidance.

My co-worker, Jorge Pena, I am deeply grateful to him for the long discussions that helped me to improve this dissertation. I am also thankful to him for providing some much needed humor and entertainment.

I am greatly appreciative of my tutor Prof. Dr. Ana Martín Marcos and the members of my doctoral committee for their assistance and accessibility.

I would like to thank to many processionals in The World Bank, who working on Investment Climate Assessments (ICAs) for providing me with datasets and further insight on this interesting topic. Particularly, I would like to acknowledge José Luís Guasch, Aurora Ferrari, Inderbir Sigh Dhingra, Maria Luisa Motta, Eric Manes and Paulo Correa.

I am also indebted to Universidad Carlos III de Madrid, and especially to department of economics, for its academic and financial support. I am also grateful to the financial support received from the Chair Telefónica-UC3M of economics of telecommunications.

This work has beneficed with useful comments received from different institutions: The World Bank, Universidad Carlos III de Madrid, UNED… Especially relevant from Ariel Pakes, James Levisohn, Eric Verhoegen, Klaus Desmet, Isabel Sánchez and Rodolfo Stucchi.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of the project.

Table of Contents

Chapter I: Summary (in Spanish) **1**

Chapter II: Manufacturing Sector **23**
1. Introduction 29
2. Data 35
3. Econometric estimation of IC elasticities and semi-elasticities on productivity (TFP) 36
4. Econometric analysis of IC and productivity impact on employment, real wages, probability of exporting and probability of receiving FDI. 45
5. IC assessment on aggregate productivity (Olley and Pakes decomposition) and other measures of economic performance. 49
6. Empirical results 53
7. Conclusions 60

References 63
Appendix A: Definitions 66
Appendix B: Tables and Figures 71

Chapter III: ICT Sector **99**
1. Introduction 105
2. Data 106
3. Econometric estimation of IC elasticities and semi-elasticities on productivity (TFP) 107
4. Econometric analysis of IC and productivity impact on employment, real wages, probability of exporting and probability of receiving FDI. 113
5. IC assessment on aggregate productivity (Olley and Pakes decomposition) and other measures of economic performance. 117
6. Empirical results 121
7. Conclusions 128

References 130
Appendix A: Definitions 133
Appendix B: Tables and Figures 138

Chapter IV: Retail Sector **171**
1. Introduction 177
2. Analysis of Labor-Productivity 181
3. Analysis of Employment Demand 186
4. Conclusions 189

References 191
Appendix A: Econometric Models 194
Appendix B: Data Transformations 202
Appendix C: Definitions 205
Appendix D: Tables and Figures 208

Chapter I:

Economic Determinants of India`s Investment Climate

Summary

Table of Contents

1. Economía de la India. Perspectiva histórica **5**

2. Encuestas del Clima de Inversión **7**

3. Metodología

3.1 Sector Manufacturero y Sector TIC 8

3.2 Sector de distribución minorista 12

4. Motivación y objetivos **15**

5. Principales resultados **17**

Referencias bibliográficas **20**

1. Economía de la India. Perspectiva histórica.

Cuando obtuvo su independencia en 1947, la India era un país con una renta per cápita muy baja y contaba con una gran parte de su población viviendo bajo el umbral de la pobreza. A comienzos del siglo XXI ha conseguido reducir de una manera significativa su tasa de pobreza y se ha situado entre los países con una renta per cápita media. En estos algo más de 60 años, los éxitos del país se han desarrollado en distintos frentes:

- Consolidación de un sistema democrático.
- Reducción de la pobreza en términos de pobreza absoluta en más de la mitad.
- Importante aumento de la alfabetización.
- Muy considerable mejora de las condiciones sanitarias.
- Ha llegado a ser una de las principales economías emergentes del mundo, con tasas de crecimiento en los primeros años del nuevo siglo superiores al 9%.
- Ha emergido como uno de los mejores competidores mundiales en sectores como las tecnologías de la comunicación o el sector farmacéutico.

Las elevadas tasas de crecimiento continuado en la India durante las últimas décadas han catapultado al país a ser una de las mayores economías del mundo. En 1980 ocupaba el lugar número 50 en dólares nominales. En la actualidad se encuentra entre las 10 mayores economías mundiales y en términos de paridad del poder adquisitivo (PPA), ocupa el cuarto lugar, sólo por detrás de Estados Unidos, Japón y Rusia.

Si bien el crecimiento de la India comienza en la década de los 70, es tras las reformas llevadas a cabo a principios de los años 90 cuando el crecimiento es mayor y más sostenido.

La crisis fiscal del año 1991 obligó a la India a pedir un crédito al FMI por valor de 1.8 billones de dólares y a acometer profundas reformas en su estructura económica. Empezando por medidas de estabilización de su balanza de pagos con una devaluación de la Rupia de aproximadamente un 25%. Estas medidas, necesarias para evitar el colapso financiero del país, eran insuficientes para cambiar un sistema productivo que no había sufrido grandes transformaciones desde el periodo colonial, caracterizado por agricultores e industrias de pequeña escala.

La insostenibilidad del sistema existente y la necesidad de una mejor integración en los mercados internacionales condujo al nuevo gobierno de Narasimha Rao, encabezado en los asuntos económicos por una figura no política como era el economista Manmohan Sing (actual primer ministro), a invertir la línea histórica de regulación e intervencionismo gubernamental y a llevar a cabo controvertidas reformas en áreas tales como: el sistema tributario, los servicios financieros y las administraciones públicas. Los principales objetivos de las reformas ejecutadas se centraron en:

- Reducir la importancia del sector público para que el sector privado ganará protagonismo como motor de la economía.
- Dar mayor confianza a la competencia y a las fuerzas de mercado como baluartes para el incremento de la eficiencia.
- Abrir la economía al comercio internacional, a la inversión externa y a la nueva tecnología extranjera.

La respuesta inicial a este paquete de reformas llevó a alcanzar tasas de crecimiento medio del 6.7% en los primeros años (92-97). En la segunda etapa de reformas, la tasa de crecimiento se moderó levemente situándose su media en el 5.7% (98-03). A principios del siglo XXI, en el periodo comprendido entre 2001 y 2007 (con tasas de crecimiento superiores al 9%), fue capaz de doblar la renta per cápita pasando de 400 a 800 dólares por habitante.

Los sectores líderes en estas dos décadas de elevado crecimiento continuado han sido el sector servicios y el industrial. Lo que ha causado un profundo vuelco en el reparto sectorial de la economía de la India. Representando el sector servicios a principios del 2008 un 53% del Producto Interior Bruto, siendo la participación de la industria y de la agricultura en el PIB del 29% y del 18% respectivamente.

Principalmente tres han sido los principales pilares sobre los que se ha sostenido este importante crecimiento: una mayor inversión, un aumento de la productividad y una más profunda integración en los mercados internacionales.

Para continuar con esta senda de crecimiento que mejore las condiciones de vida de la población y continúe con la reducción de los niveles de pobreza, son también tres los principales desafíos a los que la economía India tiene que hacer frente: creación de empleo en la economía formal, mayores aumentos de la productividad y reducir la desigualdad entre los distintos estados que conforman el país.

2. Encuestas del Clima de Inversión.

El Banco Mundial (World Development Report 2005) describe el clima de inversión como: i) el conjunto de factores locales concretos que configuran el conjunto de oportunidades e incentivos que tienen las empresas para realizar inversiones productivas, generar empleo y expandirse y ii) el ambiente institucional, regulatorio y normativo en que las empresas operan.

La serie de encuestas sobre el clima de inversión es parte de la nueva estrategia del Banco Mundial consistente en poner más énfasis en medir el papel de activos intangibles tales como instituciones o la cultura en el crecimiento de la actividad económica. Las encuestas recogen experiencias de las empresas en un rango de áreas relacionadas con la actividad económica como son; finanzas, gestión, corrupción, crimen, regulación, impuestos, relaciones laborales, resolución de conflictos, infraestructuras, marketing, suministros, calidad, tecnología y formación entre otros.[1]

El proyecto conjunto del clima de inversión pretende obtener información a nivel de empresa para un gran número de países en vías de desarrollo y utilizar estas bases de datos para comprender de una manera más precisa el efecto que el clima de inversión tiene en el desempeño de la actividad económica de las empresas.

El programa de clima de inversión del Banco Mundial ha conseguido obtener, mediante estas encuestas, información detallada en 70 países para más de 50.000 empresas. Las áreas de actuación han sido: Europa del este y Asia central, Asia del sur, África y Latino América. La mayor contribución de estas bases de datos es la combinación que ofrecen de información a nivel de empresa de variables de producción, experiencia y distintas facetas del clima de inversión.

Consideramos que resulta de interés explotar toda esta valiosa información desde una perspectiva microeconómica y así poder identificar todos aquellos factores del clima de inversión con una influencia significativa en la productividad total de los factores—y otras medidas de actividad económica—observada en la muestra de empresas incluidas en las ICs.[2]

---

[1] Durante los últimos años las encuestas del clima de inversión constituyen un instrumento clave para muchos investigadores aplicados y académicos que pretenden obtener evidencia empírica sobre una variedad de cuestiones particularmente interesantes en países emergentes y en vías de desarrollo. Algunos trabajos relevantes son: Reinikka y Svensson (1999), Bastos y Nasir (2003), Veeramani y Goldar (2004), Eifert et al. (2005), Haltiwanger y Schweiger (2005), Frazer (2005) y Fernandes y Pakes (2008).

[2] En este punto conviene aclarar que en este trabajo no buscamos estimar relaciones causales entre las variables del clima de inversión y diversas medidas de actividad económica. Implementar técnicas que permitan inferior relaciones causales, como son los métodos experimentales o cuasi-experimentales o bien los asociados al concepto de causalidad de Granger, no es posible con datos consistentes en simples secciones cruzadas, como es el caso de la base de datos ICs que derivan de procedimientos de encuestas realizadas sin un grupo de control.

Concretamente en este trabajo que analiza los determinantes económicos del clima de inversión en la India, utilizamos tres de estas bases de datos:

- Para el estudio del sector manufacturero se cuenta con una encuesta del clima de inversión (*Manufacturing Investment Climate Survey 2006*) que proporciona una base de datos de panel de empresas para los años 2002, 2003 y 2004.
- En el caso del análisis particular del sector de las tecnologías de la información y la comunicación, los datos utilizados proceden de la *ICT Investment Climate Survey 2006*, que contiene información para empresas en el año 2005/06.
- Finalmente, para llevar a cabo la identificación de los factores relevantes en el sector de distribución minorista contamos con la *Retail Investment Climate Survey 2006*, encuesta que recolecta datos para un total de empresas en el año 2005/06.

3. Metodología.

3.1 Sector Manufacturero y Sector TIC

Tanto en el análisis del sector manufacturero como en el del sector de las tecnologías de la información y de la comunicación la metodología empleada sigue la expuesta en los trabajos de Escribano y Guasch (2005 y 2008) y de Escribano et al (2008a y 2008b).

Básicamente, la metodología consta de dos pasos. El primero de ellos, la identificación de variables del clima de inversión con una asociación significativa con distintas variables que miden el comportamiento económico de las empresas; la PTF, la demanda de empleo, los salarios reales, las decisiones de exportar y de recibir inversión directa extranjera.

Identificación de los efectos del clima de inversión

La identificación de los efectos significativos del clima de inversión descansa en un sistema de ecuaciones estructural, que relaciona las decisiones de la empresa con el clima de inversión;

$$ptf_i = \alpha_P + \alpha_x x_i + \alpha_d d_i + \alpha_w w_i + \boldsymbol{\alpha}'_{CI}\mathbf{CI}_i^P + \boldsymbol{\alpha}'_D \mathbf{D}_i + (v_i^P + \varepsilon_i^P) \quad (A.1)$$

$$l_i = \gamma_L + \gamma_P ptf_i + \gamma_x x_i + \gamma_d d_i + \gamma_w w_i + \gamma'_{CI}\mathbf{CI}_i^P + \gamma'_D \mathbf{D}_i + (v_i^P + \varepsilon_i^P) \quad (A.2)$$

$$w_i = \beta_w + \beta_P ptf_i + \beta_x x_i + \beta_d d_i + \boldsymbol{\beta}'_{CI}\mathbf{CI}_i^w + \boldsymbol{\beta}'_D \mathbf{D}_i + (v_i^w + \varepsilon_i^w) \quad (A.3)$$

$$x_i = \delta_x + \delta_P ptf_i + \delta_d d_i + \boldsymbol{\delta}'_{CI}\mathbf{CI}_i^{Exp} + \boldsymbol{\delta}'_D \mathbf{D}_i + (v_i^x + \varepsilon_i^x) \quad (A.4)$$

$$d_i = \rho_d + \rho_P ptf_i + \rho_x x_i + \boldsymbol{\rho}'_{CI}\mathbf{CI}_i^d + \boldsymbol{\rho}'_D \mathbf{D}_i + (v_i^d + \varepsilon_i^d) \quad (A.5)$$

Donde ptf es productividad total de los factores, l es empleo, w son los salarios reales (las tres anteriores en logaritmos), x y d son respectivamente variables binarias (0,1) que sirven para estimar las probabilidad de exportar y de recibir inversión directa extranjera. $\mathbf{CI}^e$ (con e = p, l, w, x, d) es la matriz de variables del clima de inversión significativas en cada ecuación. Por último **D** es la matriz de dummies de sector, ciudad y tamaño y v_e and ε_e son errores aleatorios. En particular, de acuerdo a Escribano y Guasch (2005 y 2008) v_e es la parte del término de error, normalmente inobservable, y aquí aproximada por variables a nivel de empresa relacionadas con el clima de inversión y ε_e es un error aleatorio incorrelado con las variables explicativas del sistema. El conjunto de vectores de parámetros $\boldsymbol{\alpha}_{CI}$, $\boldsymbol{\gamma}_{CI}$, $\boldsymbol{\beta}_{CI}$, $\boldsymbol{\delta}_{CI}$, $\boldsymbol{\rho}_{CI}$ definen la asociación del clima de inversión con el comportamiento económico y son los que estamos interesados en estimar. La estimación del resto de parámetros del modelo es contingente a obtener parámetros consistentes de estos vectores. [3]

A su vez, la ecuación de PTF es obtenida de otro sistema de ecuaciones estructural, definido en las ecuaciones (A.6) a (A.8);

$$y_i = \alpha_l l_i + \alpha_m m_i + \alpha_k k_i + ptf_i \quad (A.6)$$

$$ptf_i = a_i + \alpha_w w_i + \boldsymbol{\alpha}'_D \mathbf{D}_i + \alpha_P + v_i^p \quad (A.7)$$

$$a_i = \alpha_x x_i + \alpha_d d_i + \boldsymbol{\alpha}'_{CI}\mathbf{CI}_i^P + \varepsilon_i^p \ . \quad (A.8)$$

Donde y es el logaritmo del output, l es el logaritmo del empleo, m es el logaritmo de los materiales intermedios, k es el logaritmo del stock de capital.

[3] Merece la pena clarificar en este punto que aquí entendemos la productividad total de los factores como un residuo; "Residuo de Solow", o aquella parte del output de las empresas que no puede expresarse como una combinación de la utilización de los factores productivos. Además ese residuo es "una caja negra", sobre cuyo contenido no hay un consenso claro dentro de la literatura académica. Existe debate sobre si el residuo puede usarse como una medida de la verdadera eficiencia técnica de una economía. Bajo este razonamiento, aunque durante todo el trabajo nos referiremos al residuo como la PTF, entendemos que en muchos casos puede ser que no se cumplan los supuestos para que pueda ser considerado como eficiencia técnica pura.

En concreto, para obtener el conjunto de coeficientes $\boldsymbol{\alpha_{CI}}$, $\boldsymbol{\gamma_{CI}}$, $\boldsymbol{\beta_{CI}}$, $\boldsymbol{\delta_{CI}}$, $\boldsymbol{\rho_{CI}}$ estimamos los sistemas (A.1) a (A.5) y (A.6) a (A.8), en dos etapas. Primero estimamos el sistema de productividad, y una vez que obtenemos resultados robustos podemos pasar a estimar el sistema completo de (A.1) a (A.8).

Evaluación de los efectos del clima de inversión en el comportamiento económico

Una vez identificados y estimados los coeficientes que marcan la asociación entre el clima de inversión y la actividad económica podemos pasar a la evaluación y valoración de la importancia relativa de cada factor del clima de inversión. Para evaluar y valorar la importancia relativa del clima de inversión en el caso de la PTF, partimos de la descomposición de la PTF agregada propuesta en Olley y Pakes (1996)

$$\mathrm{ptf}_r = \overline{\mathrm{ptf}}_r + N_r \hat{c}\mathrm{ov}(s_{r,i}, \mathrm{ptf}_{r,i}). \tag{A.9}$$

Donde ptf_r es la PTF agregada (media ponderada de la PTF, con ponderaciones dadas por las cuotas de mercado de cada empresa) para un país o región. $\overline{\mathrm{ptf}}_r$ es la media simple de la PTF, y el último término de la descomposición es la eficiencia asignativa, que nos dice si las empresas con mayor productividad son las que tienen mayores cuotas de mercado, y por tanto las que usan una mayor proporción de recursos de la economía, en cuyo caso el término sería positivo y grande, o si por el contrario son las empresas menos productivas las que usan los recursos lo que tiene implicaciones negativas en términos de la PTF agregada pues estas usarían los recursos menos eficientemente que las más productivas.

Si proyectamos los valores de la PTF de (A.7) y (A.8) en (A.9), obtenemos una forma reducida de la descomposición de Olley y Pakes en términos del clima de inversión. En otras palabras, siguiendo Escribano et al. (2008a), podemos expresar la PTF agregada como una suma ponderada de los valores medios del clima de inversión y las covarianzas entre el clima de inversión y las cuotas de mercado mediante la siguiente expresión

$$\begin{aligned} \mathrm{ptf}_r = {} & \hat{\alpha}_p + \hat{\alpha}_w \overline{w} + \hat{\alpha}_x \overline{x} + \hat{\alpha}_x \overline{d} + \hat{\alpha}_{IC} \overline{CI}{}^{P} + \hat{\alpha}_{DR} \overline{D} + \overline{\hat{u}}_i + N_r \hat{\alpha}'_w \hat{c}\mathrm{ov}(s_i, w_i) + \\ & + N_r \hat{\alpha}'_x \hat{c}\mathrm{ov}(s_i, x_i) + N_r \hat{\alpha}'_d \hat{c}\mathrm{ov}(s_i, d_i) + N_r \hat{\alpha}'_{IC} \hat{c}\mathrm{ov}(s_i, \overline{CI}_i^p) + \\ & + N_r \hat{\alpha}'_{Ds} \hat{c}\mathrm{ov}(s_i, D_j) + N_r \hat{c}\mathrm{ov}(s_i, \overline{\hat{u}}_i) \end{aligned} \tag{A.10}$$

Donde por simplicidad CI y D representan ahora escalares en vez de matrices y u=v+e. Concretamente, si eliminamos todo aquello que no sea clima de inversión y normalizando, de modo que la contribución del clima de inversión a la PTF agregada sea 100, la expresión (A.10) se ve reducida a la siguiente igualdad

$$100=(ptf_i^C)^{-1}[\hat{\alpha}_{IC}\overline{CI}^P+N_r\hat{\alpha}'_{IC}\hat{c}ov(s_i,CI_i^p)]\bullet 100 \tag{A.11}$$

En este caso d y x están incluidas en el CI, aunque no w. Por otro lado, ptf^C_i es la PTF Corregida en logaritmos, es decir la parte de la productividad a nivel de empresa que está influenciada por el clima de inversión. De modo que la PTF agregada corregida es un índice del efecto del clima de inversión en una economía. La descomposición de Olley y Pakes para la PTF agregada corregida es

$$ptf_r^C=\hat{\alpha}'_{IC}\overline{IC}^P+N_r\hat{\alpha}'_{IC}\hat{c}ov(s_i,IC_i^P) \tag{A.12}$$

Las ventajas de la PTF corregida es que podemos hacer comparaciones internacionales del clima de inversión sin tener que depender de la PTF regular, la cual puede estar influida por errores de medida o diferencias en los deflactores usados que están incorporados implícitamente en el término de productividad constante (α_P) entre establecimientos, inutilizando las medidas generales de la PTF para hacer comparaciones multilaterales. Mediante la PTF Corregida conseguimos eliminar este término de productividad constante, concentrándonos en lo que es exclusivamente clima de inversión, permitiéndonos hacer comparaciones entre ciudades y países. Por otro lado, supondremos que tanto las exportaciones y la inversión directa extranjera (d y x) están incluidas en el CI, concretamente entre el grupo de otras variables de control. Sin embargo no haremos lo mismo con el salario medio (w) que como medida del capital humano está más relacionada con otros factores distintos del clima de inversión. Por lo tanto, como ya hemos apuntado, ptf^C_i es la PTF Corregida, es decir la parte de la productividad a nivel de empresa que está influenciada por el clima de inversión (para más detalles ver Escribano et al., 2008a).

Mediante la evolución de la importancia relativa de cada factor del clima de inversión expuesta podemos; i) comparar contribuciones netas del clima de inversión a la PTF, sin tener en cuenta el impacto de los salarios reales, la constante de productividad, o dummies de sector, ciudad o tamaño; ii) podemos saber qué parte es debida a las infraestructuras, a la burocracia, etc; iii) podemos tomar las contribuciones en valor absoluto para comparar contribuciones absolutas, evitando que contribuciones negativas se compensen con contribuciones positivas; iv) finalmente, podemos comparar contribuciones por país, ciudad o incluso por tamaño de empresa.

Simétricamente, para el resto de medidas de la actividad económica realizamos una valoración similar, en este caso nos centramos en los valores medios de los salarios, el empleo, las exportaciones y la inversión directa extranjera. De un modo formal, tomamos las ecuaciones (A.2) a (A.5) en las medias y, tras sustituir los parámetros del modelo por sus correspondientes estimaciones tenemos:

$$100 = \overline{l}_i^{-1} \bullet [\hat{\gamma}_L + \hat{\gamma}_p \overline{ptf}_i + \hat{\gamma}_d \overline{d}_i + \hat{\gamma}_w \overline{w}_i + \hat{\boldsymbol{\gamma}}'_{\mathbf{CI}} \overline{\mathbf{CI}}_{\mathbf{i}}^{\mathbf{P}} + \hat{\boldsymbol{\gamma}}'_{\mathbf{D}} \overline{\mathbf{D}}_{\mathbf{i}} + (\overline{\hat{v}}_i^P + \overline{\hat{\varepsilon}}_i^P)] \bullet 100 \qquad (A.13)$$

$$100 = \overline{w}_i^{-1} \bullet [\hat{\beta}_w + \hat{\beta}_P \overline{ptf}_i + \hat{\beta}_x \overline{x}_i + \hat{\beta}_d \overline{d}_i + \hat{\boldsymbol{\beta}}'_{\mathbf{CI}} \mathbf{CI}_i^w + \hat{\boldsymbol{\beta}}'_{\mathbf{D}} \mathbf{D}_i + (\overline{\hat{v}}_i^w + \overline{\hat{\varepsilon}}_i^w)] \bullet 100 \qquad (A.14)$$

$$100 = \overline{x}_i^{-1} \bullet [\hat{\delta}_x + \hat{\delta}_P \overline{ptf}_i + \hat{\delta}_d \overline{d}_i + \hat{\boldsymbol{\delta}}'_{\mathbf{CI}} \overline{\mathbf{CI}}_i^{\mathbf{Exp}} + \hat{\boldsymbol{\delta}}'_{\mathbf{D}} \overline{\mathbf{D}} + (\overline{\hat{v}}_i^x + \overline{\hat{\varepsilon}}_i^x)] \bullet 100 \qquad (A.15)$$

$$100 = \overline{d}_i^{-1} \bullet [\hat{\rho}_d + \hat{\rho}_P \overline{ptf}_i + \hat{\rho}_x \overline{x}_i + \hat{\boldsymbol{\rho}}'_{\mathbf{CI}} \overline{\mathbf{CI}}_i^{\mathbf{d}} + \hat{\boldsymbol{\rho}}'_{\mathbf{D}} \overline{\mathbf{D}} + (\overline{\hat{v}}_i^d + \overline{\hat{\varepsilon}}_i^d)] \bullet 100 \qquad (A.16)$$

3.2 Sector de distribución minorista.

Análisis de la productividad del trabajo

En el caso concreto del estudio del sector de distribución minorista, la metodología empleada es la desarrollada en el propio trabajo para identificar los factores relevantes que afectan a la productividad del trabajo y a la demanda de empleo en las empresas pertenecientes a dicho sector y cuantificar su impacto.

El concepto de productividad utilizado aquí es el que se refiere a la productividad del trabajo, debido principalmente a la ausencia de información de variables de producción como el stock de capital. Como los objetivos siguen siendo los mismos, para identificar los efectos significativos del clima de inversión partimos de la forma reducida de la ecuación de la productividad del trabajo en términos de los precios de los inputs (w y r), de las variables del clima de inversión (IC) y de otras variables de control (C), siendo ésta,

$$\log\left(\frac{Y}{L}\right)_i = \gamma_0 + \gamma_1 \log w_i + \gamma_2 \log r_i + \gamma'_3 IC_i + \gamma'_4 C_i + \varepsilon_{i,YL} \qquad (B.1)$$

Donde el subíndice i=1,…,1987 denota la observación para cada empresa. El tratamiento de las variables del clima de inversión (IC) continúa siendo el mismo que el propuesto en Escribano y Guasch (2005 y 2008), como *efectos fijos observables*.

Bajo condiciones normales, las estimaciones por mínimos cuadrados ordinarios (MCO) de los parámetros de la ecuación B.1, son consistentes una vez que controlamos por las variables del clima de inversión.

De manera similar, podemos pasar a la evaluación y valoración de la importancia relativa de cada factor del clima de inversión. Para ello, llevamos a cabo la descomposición de la productividad agregada del trabajo en los dos componentes establecidos por Olley y Pakes en 1996, siendo su formulación:

$$\log\left(\frac{Y}{L}\right)_j = \overline{\log}\left(\frac{Y}{L}\right)_j + \sum_{i=1}^{N_j} \tilde{s}_{j,i}^{Y} \log\left(\frac{\tilde{Y}}{L}\right)_{j,i} \tag{B.2}$$

Siendo el primer término de la derecha de la ecuación la productividad del trabajo media y el segundo, la eficiencia asignativa o covarianza entre la cuota de mercado y la productividad del trabajo multiplicado por el número de empresas. La interpretación de los resultados obtenidos a partir de esta descomposición sigue los mismos criterios que en Escribano y Guasch (2005 y 2008).

Por último para cuantificar la contribución individual de cada variable del clima de inversión en la productividad del trabajo, evaluamos la ecuación B.1 en la media,

$$100 = \hat{\gamma}_0 \left(\frac{1}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_1 \left(\frac{\overline{\log}\, \overline{w}}{\overline{l\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_2 \left(\frac{\overline{\log}\, \overline{r}}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_3' \left(\frac{\overline{IC}}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_4' \left(\frac{\overline{C}}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 \tag{B.3}$$

Análisis de la demanda de trabajo

Comúnmente la demanda condicional de empleo se expresa en función del nivel de output que la empresa desea producir, de la productividad total de los factores (PTF), de los precios de los inputs (salarios, coste de uso del capital…) y de las variables del clima de inversión (IC) y otras variables de control (C)

$$\log L_{it} = \beta_0 + \beta_1 \log\left(\frac{Y}{L}\right)_{it} - \beta_2 \log w_{it} + \beta_3 \log r_{it} + \beta_4' IC_i + \beta_5' C_{it} + \varepsilon_{L,it} \tag{B.4}$$

Para estimar los coeficientes de la ecuación B.4, llevamos a cabo una regresión por mínimos cuadrados en dos etapas, ya que la productividad del trabajo es una variable endógena en la ecuación B.4.

Por último para cuantificar el impacto de las variables del clima de inversión en la demanda de trabajo, seguimos el mismo procedimiento y evaluamos la ecuación B.4 en su media, resultando la siguiente expresión:

$$100 = \hat{\beta}_0\left(\frac{1}{\overline{\log L}}\right)100 + \hat{\beta}_1\left(\frac{\overline{\log}\left(\frac{Y}{L}\right)}{\overline{\log L}}\right)100 + \hat{\beta}_2\left(\frac{\overline{\log w}}{\overline{\log L}}\right)100 + \hat{\beta}_3\left(\frac{\overline{\log r}}{\overline{\log L}}\right)100 + \hat{\beta}_4\left(\frac{\overline{IC}}{\overline{\log L}}\right)100 + \hat{\beta}_5\left(\frac{\bar{C}}{\overline{\log L}}\right)100 \qquad \text{(B.5)}$$

4. Motivación y objetivos

Las primeras versiones de estos tres trabajos forman parte del análisis econométrico que sirve como base empírica al informe publicado por el Banco Mundial "India`s Investment Climate: Voices of Indian Business", dirigido por Aurora Ferrari e Inderbir Sigh Dhingra.

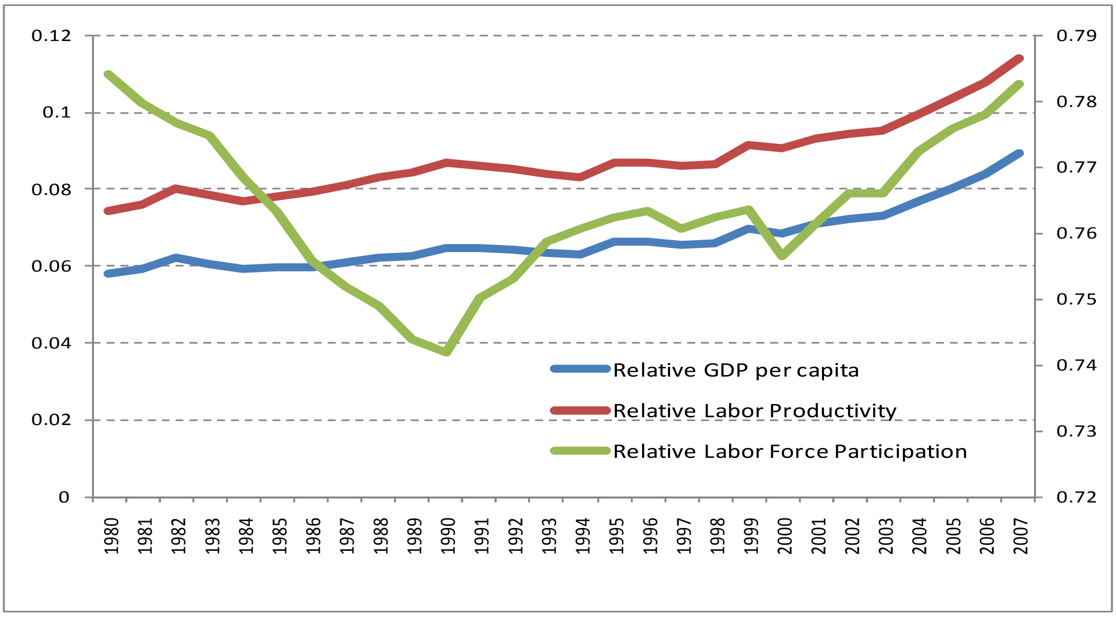
El mencionado informe trata de identificar los principales cuellos de botella, dentro del clima de inversión, que obstaculizan el crecimiento y la reducción de la pobreza en la India. Concretamente pretende dar respuesta a los tres desafíos principales a los que se enfrenta la economía del país: i) ¿cómo incrementar la productividad?, ii) ¿cómo acelerar la creación de empleo?, y iii) ¿cómo reducir las desigualdades existentes entre los distintos estados que conforman el país?

Los resultados del trabajo muestran la importancia de los retos a los que la India se enfrenta y cómo la mejora del clima de inversión en sectores estratégicos ayuda a la economía de la India a superar estos tres desafíos ya mencionados.

Desde el convencimiento en la necesidad de tal objetivo y con la certeza del papel relevante que un clima de inversión adecuado juega en el desarrollo de una economía, analizamos la evolución del PIB de la India en los últimos 30 años. Como nos ilustra la Figura

1, el crecimiento que ha experimentado la India en las últimas décadas y su convergencia con los Estados Unidos, se ha cimentado en mejoras en la productividad del trabajo.

Figura 1: Descomposición del diferencial de PIB entre la India y los Estados Unidos, 1980/2007



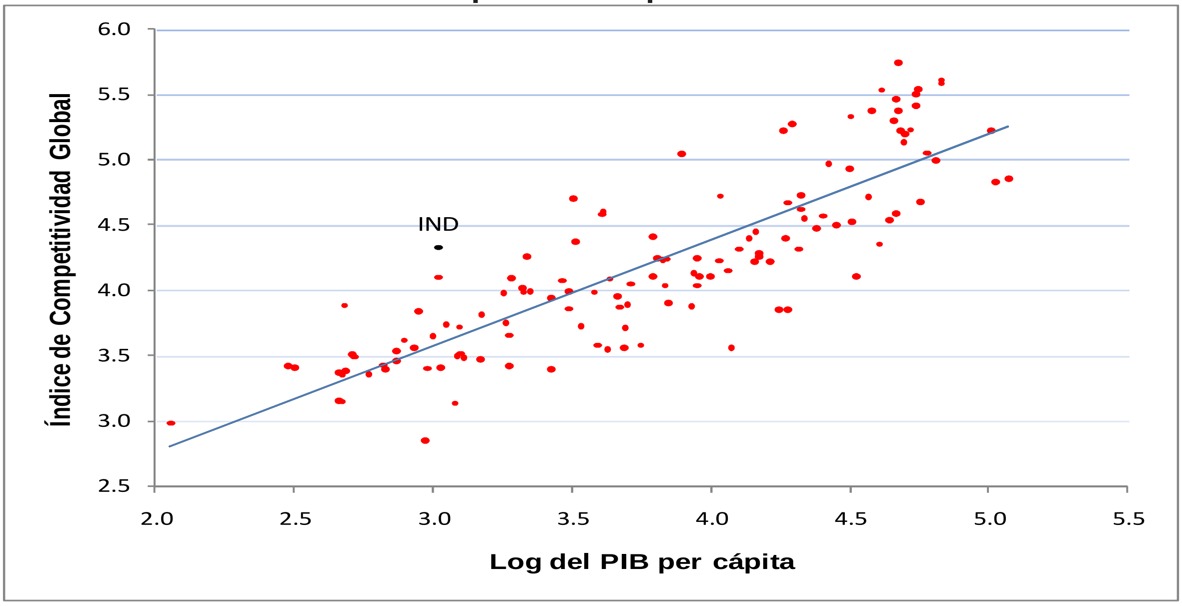
Fuente: Elaboración propia a partir de Penn World Table Version 6.3, Center for International Comparisons at the University of Pennsylvania.

Asumiendo que el crecimiento del PIB en un país lleva consigo mejoras en las condiciones de vida de sus ciudadanos y reducción de sus niveles de pobreza y que la productividad total de los factores es un componente muy importante de la productividad del trabajo, es directa la implicación que nos lleva a estudiar los determinantes que afectan a la PTF en la India. En este caso buscamos identificar y evaluar la importancia de los determinantes del clima de inversión en la productividad y en otras medidas de la actividad económica como son: la demanda de empleo, los salarios reales, la probabilidad de exportar y la probabilidad de recibir inversión directa extranjera.

5. Principales resultados

El recientemente publicado informe "*Global Competitiveness Report*" (2008) – GCR en adelante – evalúa y clasifica 138 economías mundiales en virtud a 12 pilares o fundamentos relacionados con la competitividad. El trabajo resume sus conclusiones elaborando un índice de competitividad medio para cada país basado en el comportamiento de cada economía en los 12 pilares.[4] La Figura 2 nos muestra la fuerte correlación positiva que guarda este índice con el PIB per cápita en la muestra de 138 países. La India (IND en el gráfico) no está en una situación privilegiada en lo que a competitividad se refiere. En términos de rankings, la India ocupa la posición número 80. Sus puntos fuertes descansan en el tamaño de mercado principalmente y en factores relacionados con la innovación. Por el contrario, los condicionantes que le relegan a las 80ª posición global, tienen que ver con la inestabilidad macroeconómica y con los niveles de salud y de educación primaria.

Figura 2. Nube de puntos entre el índice de competitividad global y la renta per cápita de 138 países



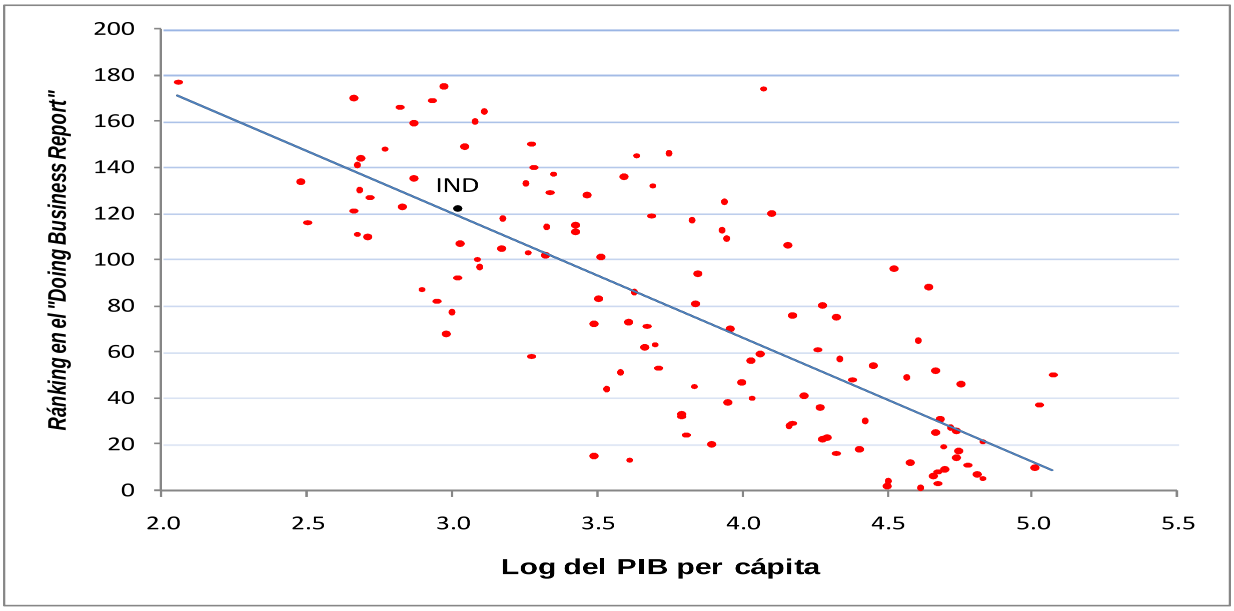
El índice se calcula como una media ponderada de los índices de los 12 pilares fundamentales de competitividad. Se tiene en cuenta el estado de desarrollo de cada economía (país) a la hora de dar los pesos a cada uno de los pilares. IND = India.
Fuente: Elaboración propia con datos de World Economic Outlook 2009, IMF; y Global Competitiveness Report 2008, The World Economic Forum.

Dar a los indicadores de la economía india una perspectiva internacional es una buena forma de ilustrar y describir sus fortalezas y sus debilidades. El informe "*Doing Business Report*" de 2009, (DBR de ahora en adelante), elaborado anualmente por el Banco Mundial en 181 economías (países), es muy ilustrativo en este sentido. En el DBR, las economías son clasificadas en función de la facilidad (ausencia de barreras) para hacer negocios; cuántos permisos o licencias son necesarios para abrir o cerrar una empresa, cuán fácil o difícil es comerciar internacionalmente o cumplir los contratos firmados, etc. Como es de esperar, las conclusiones de este trabajo para la India son consistentes con las discutidas anteriormente del informe GCR.

[4] Los 12 pilares son instituciones, infraestructuras, estabilidad macroeconómica, salud y educación primaria, educación superior y formación, eficiencia de los mercados de bienes, eficiencia mercado de trabajo, sofisticación del mercado financiero, disposición tecnológica, tamaño de mercado, sofisticación del sector negocios e innovación,

En el cómputo global, en base al ranking de facilidad de hacer negocios, India ocupa el puesto 122. La importancia de este tipo de barreras sobre la eficiencia y la competitividad queda clara al pensar en los efectos que tienen sobre los sistemas de incentivos de las empresas y en los procesos de creación destructiva.[5]

Figura 3. Nube de puntos entre el ranking sobre la facilidad de hacer negocios y la renta per cápita



El ranking final es una media ponderada de los rankings en cada uno de los aspectos básicos para hacer negocios recogidos en el DBR 2009.En el eje de ordenadas (eje vertical), un mayor número significa un mejor puesto en el ranking.
Fuente: Elaboración propia con datos World Economic Outlook 2009, IMF; Doing Business Report 2009, The World Bank Group, Washington, DC.

La importancia de tener un entorno institucional que favorezca el hacer negocios se ilustra en la Figura 3, donde la nube de puntos entre el inverso del ranking de facilidad de hacer negocios y el PIB per cápita revela una clara y esperada relación negativa entre ambas variables.

Nuestro trabajo empírico concuerda con lo que estos informes ponen de manifiesto. Se pueden identificar problemas comunes a los tres sectores en lo que se refiere al clima de inversión. Problemas con el suministro de electricidad, el sistema tributario, la corrupción y las restricciones de mano de obra cualificada son obstáculos para el incremento de la productividad y el desarrollo de la actividad económica de las empresas indias sea cual sea el sector en el que operen.

[5] Referencias ilustrando la importancia de reducir fricciones (barreras) en los mercados favoreciendo los procesos de entrada–competencia–salida con mejoras en los niveles de productividad son los trabajos de Bartelsman, Haltiwanger y Scarpetta (2004), Foster, Haltiwanger y Krizan (2006), Olley y Pakes (1996), Alfaro et al. (2007), Bartelsman, Haltiwanger y Scarpetta (2006), Hsieh y Klenow (2006) o Restuccia y Rogerson (2007).

Sector Manufacturero

Sector caracterizado por su crecimiento pero con niveles de productividad todavía bajos si lo comparamos con otros países y con otros sectores dentro de la economía de la India. La mejora de la productividad es clave para la ganancia de competitividad del sector y para el incremento de su actividad exportadora. Los empresarios identifican como principales obstáculos para el desarrollo de su actividad empresarial: el suministro eléctrico, el sistema tributario y los problemas relacionados con la corrupción. Obstáculos que pone de manifiesto nuestro análisis econométrico. Además, también resultan relevantes los factores relacionados con la calidad, la innovación y la cualificación de la mano de obra, especialmente en las empresas con mayor cuota de mercado.

Sector TIC

En sector de las tecnologías de la información y las comunicaciones es un ejemplo del éxito de las políticas de liberalización llevadas a cabo a principios de la década de los noventa. Es el sector que ha experimentado un mayor crecimiento y es el que aporta una mayor contribución a las exportaciones del país. Parte muy importante de la buena actuación del sector se debe a un clima de inversión que le proporciona ventajas en los costes. Los principales aspectos favorables de este clima de inversión son: contar con una importante mano de obra cualificada con bajo coste relativo, completa liberalización que permite transferencias tecnológicas y ventajas fiscales en las exportaciones para las empresas localizadas en parques tecnológicos de la industria del software.

Sin embargo el crecimiento potencial del sector es todavía mayor. Y existen obstáculos dentro del clima de inversión que dificultan alcanzar los niveles de productividad potenciales. Problemas con las infraestructuras y las trabas burocráticas son los principales cuellos de botella específicos que tanto las percepciones de los empresarios como nuestro análisis ponen de manifiesto.

Sector de distribución minorista

Aunque es un sector que también ha tenido crecimiento, éste ha sido pequeño comparado con el resto del sector servicios y los niveles de productividad son también pequeños si lo comparamos con el sector de distribución minorista de otros países.

Caracterizado por una exagerada densidad de establecimientos tradicionales de pequeño volumen. Las reformas introducidas, tanto por las administraciones nacionales como por las locales, han ido encaminadas a favorecer el establecimiento de tiendas con formatos modernos, más productivas. Además de los problemas comunes al conjunto de la economía, se identifica el acceso al suelo como uno de los principales obstáculos.

REFERENCIAS BIBLIOGRÁFICAS

Ahluwalia M. S. (2005). “Lessons from India`s Economic Reform”. Development Challenges in the 1990s`. *The World Bank and Oxford University Press*.

Escribano Alvaro and J. Luis Guasch (2005); “Assessing the Impact of the Investment Climate on Productivity using Firm Level Data: Methodology and the Cases of Guatemala, Honduras and Nicaragua,” *World Bank Policy Research Working Paper* # 3621, The World Bank, June.

Escribano A. and J. L. Guasch (2008); “Robust Methodology for Investment Climate Assessment on Productivity: Application to Investment Climate Surveys from Central America,” *Working Paper # 08-19 (11)*, *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena. (2008a); “Investment Climate and Firm’s Performance: Econometric and Applications to Turkey’s Investment Climate Survey,” Working Paper # 08-20 (12), *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena (2008b); “Investment Climate Assessment Based on Demeaned Olley and Pakes Decompositions: Methodology and Applications to Turkey’s Investment Climate Survey,” Working Paper # 08-20 (13), *Universidad Carlos III de Madrid*.

Wheelan, Ch. Et al (2006). “Economic Reform in India”. The Harris School of Public Policy. *The University of Chicago.*

World Bank (2009): “Doing Business Report 2009: Comparing Regulation in 181 Economies,” *The World Bank Group*, Washington DC, USA.

World Bank 2007. “Unleashing India`s Innovation: Toward Sustainable and Inclusive Growth”. Washington, D.C. World Bank

World Bank 2009. “India`s Investment Climate: Voices of Indian Business”. Washington, D.C. World Bank

World Economic Forum, The (2008).”Global Competitiveness Report”

Chapter II

Economic Determinants of India`s Investment Climate

Manufacturing Sector

Economic Determinants of India`s Investment Climate

Manufacturing Sector

Abstract

In this paper we apply the econometric methodology developed in Escribano and Guasch (2005 and 2008) and Escribano et al (2008b) for the identification of the main IC effects on economic performance. The investment climate surveys are very useful in identifying key investment climate bottlenecks that slow down growth and poverty reduction. Concretely, in this work we analyze the organized manufacturing sector, which has the second largest contribution to GDP in India (after agriculture) and is characterized for its excessive regulation and its low productivity. We have used a panel date set from 2002 to 2004, coming from Investment Climate Surveys carried out by The World Bank. The main findings of the report point to the inefficient bureaucracy and the poor infrastructures that constrain the Indian manufacturing firms economic activity.

JEL Clasification: D24; D61; J20; J30; L60; O40; O53

Keywords: *productivity, investment climate, Olley and Pakes decomposition, demeaned productivity and average contribution.*

Table of Contents

1. INTRODUCTION 29

2. DATA 35

3. ECONOMETRIC ESTIMATION OF IC ELASTICITIES AND SEMI-ELASTICITIES ON PRODUCTIVITY (TFP). 36
3.1 Robustness of IC elasticities and semi-elasticities: single step and two step estimation, restricted and unrestricted input-output elasticities. ________ 37
3.2 Endogeneity of production function (PF) variables. ________ 39
3.3 Role of prices on production function (sales generating functions and market power). ________ 39
3.4 Endogeneity of IC variables. ________ 41
3.5 Selection of the relevant models. ________ 42
3.6 Estimation under alternative replacement processes for missing data. ________ 42

4. ECONOMETRIC ANALYSIS OF IC AND PRODUCTIVITY IMPACT ON EMPLOYMENT, REAL WAGES, PROBABILITY OF EXPORTING AND PROBABILITY OF RECEIVING FDI 45
4.1 Identification of the system of equations ________ 48

5. IC ASSESSMENT ON AGGREGATE PRODUCTIVITY (OLLEY AND PAKES DECOMPOSITION) AND OTHER MEASURES OF ECONOMIC PERFORMANCE. 49
5.1 O&P decompositions: in levels and mixed. ________ 49
5.2 IC effects on productivity measure in the terms of the mixed O&P decomposition. ________ 50
5.3 Simulations based on the IC effects on the O&P decomposition of TFP. ________ 51
5.4 International comparisons of IC effects on aggregate demeaned productivity. ________ 51
5.5 IC evaluation on the sample means of employment and wages, on the probability of exporting and on the probability of receiving FDI. ________ 52

6. EMPIRICAL RESULTS 53
6.1 Key Results on Productivity ________ 53
6.2 Key Results on Employment ________ 55
6.3 Key Results on Real Wages ________ 56
6.4 Key Results on the Probability of Exporting ________ 57
6.5 Key Results on the Probability of Receiving Foreign Direct Investment ________ 58
6.7 Key Results on the Manager's Perceptions ________ 59
6.8 International Comparison. ________ 59

7. CONCLUSIONS 60

REFERENCES 63

APPENDIX A: DEFINITIONS 66

APPENDIX B: TABLES AND FIGURES 71

1. Introduction

Investigating the causes and consequences of the economic success has captured the interest of both economists and policy makers in their quest for the solution of poverty and growth and to create and help implement improved strategies for economic development.

As part of a new strategy of the World Bank on putting more emphasis on intangible assets of countries such as knowledge, institutions and culture the Investment Climate Assessments have born. As new data become available a new frontier of research is now opening to scholars. One of the new frontiers of economics is the analysis of growth from a microeconomic perspective, creating opportunities for people to escape from poverty and improve their living standards.

A significant component of country competitiveness is having a good investment climate or business environment. The investment climate, as defined in the WDR (2005), is "the set of location-specific factors shaping the opportunities and incentives for firms to invest productively, create jobs and expand." It is now well accepted and documented, conceptually and empirically, that the scope and nature of regulations on economic activity and factor markets - the so-called investment climate and business environment - can significantly and adversely impact productivity, growth and economic activity (see Bosworth and Collins, 2003; Rodrik and Subramanian, 2004; Loayza, Oviedo and Serven, 2004; McMillan, 1998 and 2004; OECD, 2001; Wilkinson, 2001; Alexander et al., 2004; Djankov et al., 2002; Haltiwanger, 2002; He et al., 2003; World Bank, 2003; and World Bank, 2004 a,b). Prescott (1998) argues that to understand large international income differences, it is necessary to explain differences in productivity (TFP). His main candidate to explain those gaps is the resistance to the adoption of new technologies and to the efficient use of current operating technologies, which in turn are conditioned by the institutional and policy arrangements a society employs (investment climate variables). Recently, Cole et al. (2004) also have argued that Latin America has not replicated Western economic success due to the productivity (TFP) gap. They point to competitive barriers (investment climate constraints) as the promising channels for understanding the low productivity observed in Latin American countries.

Government policies and behavior exert a strong influence on the investment climate through their impact on costs, risks and barriers to competition. Key factors affecting the investment climate through their impact on costs are: corruption, taxes, the regulatory burden and extent of red tape in general, factor markets (labor, intermediate materials and capital), the quality of infrastructure, technological and innovation support, and the availability and cost of finance.

For example, Kasper (2002) shows that poorly understood “state paternalism” has usually created unjustified barriers to entrepreneurial activity, resulting in poor growth and a stifling environment. Kerr (2002) shows that a quagmire of regulation, which is all too common, is a massive deterrent to investment and economic growth. As a case in point, McMillan (1988) argues that obtrusive government regulation before 1984 was the key issue in New Zealand’s slide in the world per-capita income rankings. Hernando de Soto (2002) describes one key adverse effect of significant business regulation and weak property rights: with costly firm regulations, fewer firms choose to register and more become informal. Also, if there are high transaction costs involved in registering property, assets are less likely to be officially recorded, and therefore cannot be used as collateral to obtain loans, thereby becoming “dead” capital.

Likewise, poor infrastructure and limited transport and trade services increase logistics costs, rendering otherwise competitive products uncompetitive, as well as limiting rural production and people’s access to markets, which adversely affects poverty and economic activity (Guasch 2004).

The pursuit of greater competitiveness and a better investment climate is leading countries -often assisted by multilaterals such as the World Bank - to undertake their own studies to identify the principal bottlenecks in terms of competitiveness and the investment climate, and evaluate the impact these have, to set priorities for intervention and reform. The most common instrument used has been firm-level surveys, known as Investment Climate Assessments (ICAs), from which both subjective evaluations of obstacles and objective hard-data numbers with direct links to costs and productivity are elicited and imputed. Such surveys collect data at firm level on the following themes: a) infrastructure, b) red tape, informality and others, c) finance and corporate governance, d) quality, innovation and labor skills and d) other control variables like capacity utilization, age and size of the firm, etc.

While the Investment Climate Assessments are quite useful in identifying major issues and bottlenecks as perceived by firms, the data collected is also meant to provide the basic information for an econometric assessment of the impact or contribution of the investment climate (IC) variables on productivity. In turn, that quantified impact is used in the advocacy for, and design of, investment-climate reform. Yet providing reliable and robust estimates of productivity estimates of the IC variables from the surveys is not a straightforward task since; first, the surveys do not provide panel-type data on IC variables; second, neither the production function parameters nor the functional form are observed; and third, there is an identification issue separating total factor productivity (TFP) component from the inputs of the production function.

When any of the production function inputs is influenced by common causes affecting productivity, like IC variables or other plant characteristics, there is a simultaneous equation problem. In general, one should expect the productivity to be correlated with the production function inputs (technological progress is not Hicks neutral) and, therefore, inputs should be treated as endogenous regressors when estimating production functions. This property has demanded special care with the econometric specification when estimating those productivity effects and in the choice of the most appropriate way of measuring productivity.

There is an extensive literature discussing the advantages and disadvantages of using different statistical estimation techniques and/or growth accounting (index number) techniques to estimate productivity or Total Factor Productivity (TFP). For overviews of different productivity concepts and aggregation alternatives see, for example, Solow (1957), Hall (1990), Foster, Haltiwanger and Krizan (1998), Batelsman and Doms (2000), Hulten (2001), Diewert and Nakamura (2002), Jorgenson (2003), Jorgenson, Gollop and Fraumeni (1987), Olley and Pakes (1996) and Barro and Sala-i-Martin (2004). In this paper we discuss the applicability of some of these techniques to the problem at hand and present adaptations and adjustments that provide a best fit for the described objective: estimating the productivity impact of IC variables collected through a firm-level survey (international longitudinal micro-level data sets).

Apart from the differences on the measures of the economic performance and other differences in the construction of the investment climate indicators, is the ways we approach the robustness issue what really differentiate this methodology from other econometric studies. The underlying idea is based on the fact that since there is no single salient measure of productivity, to get reliable empirical IC elasticities for policy analysis, we should provide robust empirical results using several productivity measures.

We believe that improving the investment climate (IC) is a key policy instrument to promote economic growth and to mitigate the institutional, legal, economic and social factors that are constraining the convergence of per capita income and labor productivity of India relative to more developed countries. For that, we need to identify the main investment climate variables that affect economic performance measures like total factor productivity, employment, wages, exports and foreign direct investment and this is the main goal of this paper.

The recent trade literature has emphasized the importance of firm heterogeneity in understanding export behavior. Traditional trade theory either has all firms or none of the firms in a given sector export. However, micro-level evidence shows this picture to be seriously flawed. Even within so-called export sectors, a substantial fraction of firms exclusively sell in the domestic market. Bernard and Jensen (1995, 1999), Clerides, Lach and Tybout (1998), and Aw, Chung and Roberts (2000) all find that larger and more productive firms are more likely to export. This heterogeneity shows up both across and within sectors. Moreover, these stylized facts seem to be common to both developed and developing countries. The work of Bernard and Jensen (1995, 1999), for instance, focuses on the U.S., whereas Clerides, Lach and Tybout (1998) analyze Colombia, Mexico and Morocco. The results presented in this paper on India confirm many of these stylized facts. In particular, productivity is shown to have an important impact on a firm's probability to export and larger firms are more productive. This result holds up even after controlling for a large variety of investment climate variables.

These stylized facts have given rise to a number of important theoretical contributions. Melitz (2003) proposes a monopolistic competition model with heterogeneous firms. Each firm draws its productivity from a distribution. To enter the export market, firms need to pay a fixed cost. As a result, only the larger or more productive firms will choose to export, while the smaller or less productive firms will decide to only serve the domestic market. Yeaple (2005) is able to obtain the same qualitative results, without assuming that firms are randomly assigned their productivity levels. Instead, ex ante homogeneous firms get to choose between competing technologies, and can hire workers of heterogeneous skill. Different workers have comparative advantage in different technologies. As in Melitz (2003), there is a fixed cost in accessing export markets. The model generates ex post heterogeneous firms, with the low productivity firms serving the domestic markets, and the high productivity firms exporting.

What keeps low productivity firms from exporting in both Melitz (2003) and Yeaple (2005) is the existence of a fixed cost to enter export markets. There is empirical evidence supporting this view. Das, Roberts and Tybout (2006), for instance, estimate that Colombian chemical plants need to pay a fixed cost of around $1 million to enter export markets. Other papers, such as Bernard and Jensen (2004) for the U.S. and Bernard and Wagner (2001) for Germany further substantiate the existence of fixed costs involved with exporting.

In our study on India we find, for instance, that having *fixed costs* like e-mail, web page R&D activities, security costs, having their own generator, etc., increase the probability to export. In contrast to Melitz (2003), the theoretical work by Bernard, Eaton, Jensen and Kortum (2003) suggests that fixed export costs are not needed to match the heterogeneity in export performance. They propose a model with Bertrand competition, where the price a firm can charge is bound by potential rivals. In this setup it is easier for a firm to sell at home than abroad. To export, a firm needs to overcome the hurdle of transportation costs, whereas to sell in the domestic market, transportation costs reduce the threat of foreign rivals. Therefore, firms that export will be more productive, as occurs in India.

Although much of the empirical evidence points to more productive firms becoming exporters and not the other way around (see, e.g., by Bernard and Jensen, 1999, and Clerides et al., 1998), the theory on the relation between productivity and exports is not exempt from reverse causality or the simultaneity found in India. Whereas Melitz (2003) and Bernard et al. (2003) argue that high productivity firms self select to become exporters, it is also true that access to export markets may make firms more productive. In the work by Grossman and Helpman (1991), for instance, an increase in the market size allows for more varieties being produced, thus improving the productivity of final good producers. Holmes and Schmitz (2001) propose a quality ladder model, in which entrepreneurs can use their time to either block the innovation of their rivals or to innovate and move up the ladder. They show how trade shifts the relative returns from unproductive blocking towards productive innovation. Desmet and Parente (2006) emphasize yet another mechanism: they argue that access to larger markets increases the elasticity of demand, thus increasing the incentive for firms to adopt more productive technologies.

The conventional wisdom associates foreign direct investment with higher productivity. According to Markusen (1995), one important stylized fact is that multinationals are prevalent in firms and industries with high levels of R&D, a large share of professional and technical workers, and products that are new and/or technically complex. This is in line with Dunning (1993) who argues that to overcome local barriers, multinationals must have some intangible assets, such as superior technologies or more advanced management techniques and those arguments support our empirical findings in India. Markusen (1995) refers to this as knowledge-based assets.

However, the statistical contemporaneous correlation (simultaneity) between foreign ownership and productivity does not settle the question of causality. Do foreign firms, through technology transfers, improve the productivity of the firms they acquire? Or do foreign investors select more productive firms to acquire? To use the words of Evenett and Voicu (2002), are foreign investors picking winners or creating them? In order to answer this causality questions we need to have either a control group of firms or a dynamic panel of IC variables and therefore are out of the scope of this paper.

In the case of developing countries, inward FDI may increase productivity, simply because foreign investors, often based in more advanced economies, dispose of more productive technologies. In this case, domestically owned and foreign owned firms get their productivity from different exogenous distributions. However, the positive contemporaneous correlation between foreign ownership and productivity also holds up when one focuses on FDI between developed countries. The recent theoretical work of Helpman, Melitz and Yeaple (2004) proposes a mechanism, similar to the one in Melitz (2003) that rationalizes this fact. Because of the fixed costs involved in setting up an affiliate plant abroad, only the most productive firm are able to become multinationals. Even if home firms and foreign firms get their productivity assigned from the same exogenous distribution, only the more productive foreign firms will choose to set up affiliates in the home country. This self selection issue gives rise to an endogenous difference in the productivity distribution of domestically owned and foreign owned firms.

Although these theories suggest that foreign investors would tend to improve the productivity of the firms they acquire, recent work on FDI in developed countries suggests that selection bias may be a problem. This supports the view that foreign investors may be “picking winners”. For instance, Harris and Robinson (2003) find that in the case of the UK foreign firms acquire better performing local firms, without further improving productivity after acquisition. Benfratello and Sembenelli (2006) come to a similar conclusion in the case of Italy. Other studies continue to find a positive effect from foreign ownership though. Conyon et al. (2002), for example, estimate that UK firms get a 14% productivity boost after being acquired by foreign firms.

Studies of foreign acquisitions in developing countries suggest self selection bias is less of an issue. In the case of the Czech Republic, Djankov and Hoekman (2000) and Evenett and Voicu (2002) both find evidence of technology transfers by foreign owners. Moreover, the positive impact is larger in foreign owned firms than in joint ventures. In a recent study of Indonesian manufacturing plants, Jens and Smarzynska (2005) use propensity score matching to determine what would have happened to a domestic firm had it not been acquired? They find a strong positive effect of foreign ownership. The increase in plant productivity is estimated to reach 34% three years after acquisition.

In this work on India we find that productivity is one of the main variables affecting foreign investors acquiring local firms but as in the work by Jens et al. (2005), other characteristics, such as infrastructures, innovation (foreign license), exports size of the firm and labor skills (external training and experience of the manager) also matter. However, those firms that receive foreign direct investment are not more productive after controlling for R&D activities and human capital.

Productivity has also a positive and important effect on wages. These are good news since improvements in productivity (TFP) are transformed in increases in wages. The elasticity is 0.398 meaning that a one percent increases in TFP creates a 0.398% increase in wages.

The structure of this paper is the following: section 2 provides a brief description of the data base of India. In particular, the investment climate survey (ICs) survey was done at the plant level in the manufacturing sector of India. In particular, we estimate the impact of investment climate (IC) variables and other firm control (C) characteristics on several economic performance variables. (See Table A.1 of the appendix A). The properties and quality of the observations are analyzed in Tables B.I to B.III of appendix B. The IC variables are groped in five broad categories: a) infrastructure, b) red tape, informality and others, c) quality, innovation and labor skills, d) finance and corporate governance and e) other firm control characteristics. (See Table A.I to A.V of the appendix A). Sections 3, 4 and 5 present the econometric methodology. The empirical results are explained in section 6.

Notice that we also computed the Olley and Pakes (1996) decomposition of aggregate productivity in: average productivity and the efficiency term. The impact of IC variables on the efficiency component of the Olley and Pakes (1996) decomposition is also analyzed. All those decompositions of aggregate productivity are performed at several levels of aggregation: aggregate level (whole manufacturing sector), by industry, by region, by size of the firm and by year. The impact of IC variables on: average (log) productivity, average (log) wages, average (log) employment, on the probability of exporting and on the probability of receiving foreign direct investment. In section 6 we compare India´s performance with ten selected economies. Finally, section 7 includes some conclusions.

2. Data

The World Bank Group in close partnership with public or private institutions in each country creates firm level data for investment climate assessments. The surveys of private enterprises try to find out about the difficulties that firms encounter in starting and running a business—and, if the business fails, in exiting. The ICs survey captures firms' experience in a range of areas—financing, governance, regulation, tax policy, labor relations, conflict resolution, infrastructure services, supplies and marketing, technology, and training.

The survey covers formal manufacturing firms across 16 Indian states and 8 industries (Table A1). From India's IC survey we are able to form a balanced panel data base. The panel is short in the time dimension with 3 years of observations but, long in the cross section dimension. About the input variables of the production function we have temporal observations for the years 2002, 2003 and 2004. However, for the long list IC variables included in Table A2 (I-III) of the appendix, we only have observations for the year 2004. In the empirical application we assume that the investment climate characteristics for this short period of time (say three years) are constant at the plant level and therefore we treat them as observable fixed effects. This assumption has important econometric advantages, as will become clear later on.

We do not estimate the productivity equations in first differences since we will lose all the information on ICA variables, since that information is fixed (constant). In particular, we estimate the elasticities and semi-elasticities of the ICA and other control variables based on productivity measures in levels (logs), adding always dummy variables to control for the three years and the eight sectors including industries, services, farm-fishing. After appropriate handling of outliers and missing observations we end up losing several observations. Tables B.I list the number of missing observations we encounter in the IC survey of India´s manufacturing sector. Tables B.II and B.III show the representativeness of production function variables before and after cleaning for outliers and missing observations. Notice, we were able to save many IC variables by substituting the missing values by region-industry averages instead of their individual firm observation or by substituting the whole IC variable by their corresponding region industry averages. As will become clear later on, this region-industry transformation helps us also reducing the degree of endogeneity of IC variables.

3. Econometric estimation of IC elasticities and semi-elasticities on productivity (TFP).

In the identification of the significant investment climate effects on economic performance (productivity, demand for labor, real wages, probability of exporting and probability of receiving FDI) it is important to condition on the whole set of information contained in the IC survey. In particular, we propose a simultaneous equations system that relates the interactions between the investment climate variables and firm's economic performance measures.

Escribano and Guasch (2005, 2008), model that relates IC and C variables with firm-level productivity (TFP) **by the following system of equations with fixed effects**,

$$\log Y_{it} = \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \log TFP_{it} \quad (3.1a)$$

$$\log TFP_{it} = a_i + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + w_{it} \quad (3.1b)$$

$$a_i = \alpha'_{IC} IC_{P,i} + \alpha'_C C_{P,i} + \varepsilon_i \quad (3.1c)$$

where, *Y* is firms' output (sales), *L* is employment, *M* denotes intermediate materials, *K* is the capital stock, *IC* and *C* are time-fixed effect vectors of other investment climate and control time-fixed effects, and D_r, D_j and D_t are the vectors of state, industry and year dummies.

The usually unobserved time fixed effects (a_i) of the TFP equation (3.1b) are here proxy by the set of observed time fixed components IC, and C variables of (3.1c) and a remaining unobserved random effects (ε_i). The two random error terms of the system, ε_i and w_{it}, are assumed to be conditionally uncorrelated with the explanatory L, M, K, IC and C variables[6] of equation (3.2),

$$\log Y_{it} = \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \alpha'_{IC} IC_{P,i} + \alpha'_C C_{P,i} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + u_{it} \quad (3.2)$$

Therefore, the regression equation (3.2) represents the *conditional expectation* plus a composite random-effect error term equal to $u_{it} = \varepsilon_i + w_{it}$.

Before introducing the remaining equations of the system we explain the main econometric issues that we have to address in the estimation of productivity (TFP) equations.

3.1 Robustness of IC elasticities and semi-elasticities: single step and two step estimation, restricted and unrestricted input-output elasticities.

By simply plugging (2.1c) into (2.1b) we get the next expression for productivity

$$\log TFP_{it} = \alpha'_{IC} IC_i + \alpha'_C C_i + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + u_{it} \quad (3.3)$$

where *IC* and *C* are, respectively, the observable fixed effects vectors of investment climate and control variables listed in Tables 2.1 to 2.7 of the Appendix. In the regressions, we always control for several region dummies (D_r, r= 1, 2… R), sector-industry dummies (D_j, j = 1, 2... q_D), a constant term (α_P) and in the panel data case we also include a set of time dummies (D_t, t = 1, 2... q_T). Since there is no single salient measure of productivity (or *logTFP*$_{it}$), any empirical evaluation of the productivity impact the IC might critically depend on the particular productivity measure used. Escribano and Guasch (2005, 2008) suggested–following the literature on

[6] Under this formulation (and other standard conditions) the OLS estimator of the productivity equation (3.2) with robust standard errors is consistent, although a more efficient estimator (GLS) is given by the random effects (RE) estimator that takes into consideration the particular covariance structure of the error term, $\varepsilon_i + w_{it}$, which introduces a particular type of heteroskedasticity in the regression errors of (3.2).

sensitivity analysis of Magnus and Vasnev (2006)–to look for empirical results (elasticities) that are *robust to several productivity measures*. This is also the approach we follow in this paper.

In particular, we want the elasticities of IC on productivity (TFP) to be robust (with equal signs and similar magnitudes) for the 10 different productivity measures used. The alternative productivity measures used come from considering:

a) different functional forms of the production functions (Cobb-Douglas and Translog),

b) different sets of assumptions (technology and market conditions) to get consistent estimators based on Solow's residuals, ordinary least squares (OLS), or random effects (RE), and so on,

c) different aggregation levels when measuring input-output elasticities (industry level or aggregate country level).

Table 1: Summary of Productivity Measures and Estimated Investment Climate (IC) Elasticities

1. Solow Residual	Two Step Estimation	1.1 Restricted Coef. 1.2 Unrestricted Coef.	1.1.a OLS 1.1.b RE 1.2.a OLS 1.2.b RE	2 (P_{it}) measures 4 (IC) elasticities
2. Cobb-Douglas	**Single Step Estimation**	**2.1 Restricted Coef.** **2.2 Unrestricted Coef.**	**2.1.a OLS** **2.1.b RE** **2.2.a OLS** **2.2.b RE**	**4 (P_{it}) measures** **4 (IC) elasticities**
3. Translog	**Single Step Estimation**	**3.1 Restricted Coef.** **3.2 Unrestricted Coef.**	**3.1.a OLS** **3.1.b RE** **3.2.a OLS** **3.2.b RE**	**4 (P_{it}) measures** **4 (IC) elasticities**
Total				**10 (P_{it}) measures** **12 (IC) elasticities**

Restricted Coefficients= Equal input-output elasticities in all industries. Unrestricted Coefficients = Different input output elasticities by industry.
OLS = Pooling Ordinary Least Squares estimation (with robust standard errors). RE = Random Effects estimation.

Table 1 above summarizes the productivity measures used for the IC robust evaluation. The two-step estimation starts from the nonparametric approach based on cost shares from Hall (1990) to obtain Solow's residuals in logs under two different assumptions:[7] (a) the cost shares are constant for all plants located in the same country (restricted Solow residual), and (b) the cost shares vary among industries in the same country (unrestricted by industry Solow residual). Once we have estimated the Solow residuals ($logTFP_{it}$) in the first step, in the second step we can estimate equation (3.3) by OLS with robust standard errors and allowing for

[7] The advantage of the Solow residuals is that they require neither the inputs (L, M, K) to be exogenous nor the input-output elasticities to be constant or homogeneous (Escribano and Guasch, 2005 and 2008). The drawback is that they require having constant returns to scale (CRS) and, at least, competitive input markets.

clustering correlation within industries and states. For further robustness we use the available panel data for productivity and production function variables and estimate (3.3) also by RE.

In the single-step estimation approach, we start with the OLS parametric estimation (and RE for the case of the panel) of the extended production function (3.2). We use two different functional forms of the PF—Cobb-Douglas and Translog—under two different aggregation conditions on the input-output elasticities: equal input-output elasticities in all industries (restricted case) and different input-output elasticities by industries (unrestricted case).

3.2 Endogeneity of production function (PF) variables.

There is an identification issue separating TFP from PF when any PF inputs is influenced by unobserved common causes affecting productivity—such as a firm's fixed effects. This creates simultaneous equation bias if least squares are used estimating equation (3.1a) to measure TFP. However, this endogeneity problem of the inputs is overcome by using the single step least squares estimation of equation (3.2) follow the approach proposed by Escribano and Guasch (2005, 2008). That is, in (3.2) we proxy the usually unobserved firm-specific fixed effects (which are the main cause of inputs' endogeneity) by a long list of observed firm-specific fixed effects coming from the investment climate surveys. Controlling for the largest set of IC variables and plant C characteristics, we can—under standard regularity conditions— get consistent and unbiased least squares estimators of the parameters of the PF and the corresponding IC elasticities on TFP in one step.

Notice that even if we were only interested in assessing the impact of one block of IC variables, say infrastructure, we do not limit the scope of the analysis to only that block of IC variables. We include (and therefore control for) IC factors from all the blocks because of the crucial role IC variables play as proxies for the unobserved fixed effects. This is the key feature of the Escribano and Guasch (2005, 2008) econometric methodology to provide robust empirical regularities. If for example, we try to estimate the impact of say infrastructure, without controlling for the other IC blocks of variables, we can get different signs on certain coefficients due to the omitted variables problem; see Escribano and Guasch (2008).

3.3 Role of prices on production function (sales generating functions and market power).

The role of prices in the system (3.1a)-(3.1c) deserves special attention. As our dependent variable is sales, rather than units of physical output, it reflects prices. In fact, according to the current literature, the term *sales generating function* seems more appropriate rather than production function for equation (3.1a), as in the work of Olley and Pakes (1996). If prices are not identical across firms, what seems to be a high productive plant may be just an establishment that is charging high prices, what in turn may be consequence of either market power (non zero mark-ups) or differences in quality of final goods. While with homogeneous products high productivity could be a reflection of high prices, or in other words a reflection of market power (Melitz, 2000; Bernard, et al., 2003; Katayama, et al., 2006; Foster et al, 2008), under heterogeneous or differentiated products high prices could be consequence of higher quality, what could be translated to over-measured productivity as some plants would be able to produce higher quality—and price—products with the same amount of output (Levinsohn and Melitz, 2002; de Loecker, 2007; Katayama, et al., 2006; Gorodnichenko, 2007). These points are especially important in developing countries where usually market power is a severe constraint to growth. Addressing these issues is not a straightforward task with the data available. A more comprehensive analysis would need information on plant level input prices to incorporate the demand side of the model.

As long as this data is not available a plausible solution is to estimate the system (3.1a)-(3.1c) by following a control approach. Now instead of observing output (Y) we are observing sales (P_yY), where P_y denotes prices, and then equation (3.1a) is transformed to (3.1a')

$$\log Y_{it} + \log P_{y,it} = \log P_{y,it} + \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \log TFP_{it} \quad (3.1a')$$

Notice that as long as we control for *logPy* on the right hand side of equation (3.1a'), *productivity* in the RHS of the equation still is *logTFP*. Since, within a year there is low price variability at the firm level we assume that $logP_y$ can be proxied by a constant term, control variables *C* that are time-firm level fixed effect vectors of firm variables and a set of dummy variables, and D_r, D_j and D_t including the vectors of state, industry and year dummies. Therefore, after including all those variables we could assume that that $\log P_{y,it} \approx \alpha'_C C_{P,i} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t$ and therefore we can get a similar expression for (3.2) incorporating prices

$$\log Y_{it} + \log P_{y,it} = \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \alpha'_{IC} IC_{P,i} + \alpha'_C C_{P,i} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + u_{it} \quad (3.2')$$

Estimating sales in (3.2'), as we do in our empirical analysis, can provide evidence that TFP can be "interpreted" as "technical efficiency".[8] Finally, to control for the mark up (market power effect) and/or quality (differentiated products) we are also including several IC and C variables

[8] Notice, however, that the word technical efficiency that you use is too narrow in the ICs context since there are many efficiencies related to IC variables on TFP that are not technical (regulatory, governance, institutional, etc.).

related to competition (see the list of IC variables included in the group of other control variables).

3.4 Endogeneity of IC variables.

Another econometric problem we have to face when estimating the parameters of IC, and C variables—either from the two-step or single-step procedure—is the possible endogeneity of some of these explanatory variables. That is, many IC variables are likely to be determined simultaneously along with any TFP measure. With these productivity equations, the traditional instrumental variable (IV) approach is difficult to implement, given that we only have information for one year, and therefore we cannot use natural instruments, such as those provided by their own lags. As an alternative correction for the endogeneity of the IC variables, we use the region-industry-size average of plant-level IC variables instead of the crude IC variables,[9] which is a common solution in panel data studies at the firm level[10].

However, one should avoid including too many industry-region-size variables since it may lead to multicollinariety problems. Especially, if the number of states, sizes and industries is not large enough and there are common regions and/or industries processes affecting the variables. So a proper *a priori* consideration of the endogeneity of IC and C variables is important.

Using industry-region-size averages also mitigates the effect of having certain missing individual IC observations at the plant level, which—as mentioned in Section 2—represent one of the most important difficulties using ICSs. As an alternative, we also follow a second strategy to deal with the missing values of some IC, and C variables. In order to keep as many observations in the regressions as possible to avoid losing efficiency, when the response rate of the variables is large enough, we decided to replace those missing observations with the corresponding industry-region-size average.[11] Thus, we gain observations, efficiency, and representativity maybe at the cost of introducing some measurement errors into the explanatory variables.[12]

For those variables which endogeneity is intrinsic due to the construction of the simultaneous system of equations (exporting probability and probability of receiving FDI inflows)

[9] For the creation of cells a minimum number of firms are imposed—there must be at least 15 to 20 firms in each industry-region-size cell to create the average, otherwise we apply the region-industry averages. If the problem persists, we apply the industry-size or the region-size average.

[10] This two-step estimation approach is a simplified version of an instrumental variable estimator (two-stage least squares, 2SLS).

[11] Notice that this replacement strategy has a straightforward weighted least squares interpretation since we are giving a greater weight to those observations with more variance (Escribano at al., 2008b).

[12] Depending on the assumption we make, the measurement error may introduce a downward bias in the parameters that depends on the ratio between the variances of the variables and the measurement error. Since those explanatory variables are constant within regions, sizes, and industries we expect their variances will be small.

we apply standard IV estimators (2SLS) using as instruments either the industry-region-size average or those exogenous IC variables from the list of explanatory covariates of the corresponding equation.

Unfortunately, endogeneity is yet an unsettled issue in ICSs. Implementation of those techniques that allow obtaining causal interpretations, like those derived from the concept of '*Granger causality*' or experimental or quasi-experimental methods, are unfeasible to implement in the actual context of IC surveys with cross-sectional dataset or with incomplete panels with a very short time dimension. Although the solutions proposed to deal with endogeneity in this report can reduce the degree of endogeneity of both IC and PF variables, they do not allow us to place causal interpretations on the results obtained. Rather, we have to satisfy ourselves by obtaining empirical regularities with the relationships among IC variables and measures of firms' economic performance.

3.5 Selection of the relevant models.

The econometric methodology applied for the selection of the variables (IC, and C) goes from the general to the specific. The otherwise omitted variables problem that we encounter—starting from a too-simple model—generates biased and inconsistent parameter estimates. We start the selection of IC variables with a wide set compounded by up to 110 variables. We avoid using at the same time in the regression, explanatory IC variables that provide similar information (highly correlated), mitigating the problem of multicollinearity that could otherwise arise. We then start removing from the regressions—the less significant variables—one by one, until we obtain the final set of IC variables, significant in at least one of the alternative TFP regressions and with parameters varying within a reasonable range of values. Once we have selected a preliminary model we test for omitted IC variables (those initially dropped IC variables).

The robust TFP effects obtained on IC and C variables, along with their level of significance, are listed in Table C.I and Table C.II of the appendix B included at the end of the report. Indications of the form the variables are entering the regression—industry-region-size average or missing values replaced by the industry-region-size average, logs, etc.—are also included in the Table. In all the cases we are using robust standard errors.

3.6 Estimation under alternative replacement processes for missing data.

Following Escribano and Pena (2009), a second set of robust results is applied by using different imputation mechanisms for replacing missing values in production function figures. We allow for different assumptions on the missing data mechanism (MDM): missing at random (MAR) or non-ignorable MDM, for more information on MDM see for instance Little and Rubin (1987). As we have already pointed out in section two of this paper, the simple imputation mechanism that we usually applied in ICSs (called ICA method) is based on the conditional expectation of each of the missing production function variables on firm's information by industry, region and size. Here we analyze their robustness under alternative imputation methods, briefly described below; for a complete explanation of these imputation methods see Escribano and Pena (2009).

We first describe alternative replacement or imputation mechanism to the ICA method that rely on the missing at random (MAR) assumption:

a) **Bootstrap ICA method:** If we use imputed observations as if they were real data the resulting regression standard errors estimates will be in general too low and inference might lead to find too many significant variables. This has to do with the lack of uncertainty in the estimation of the parameters of the model. Conventional formulas to compute standard errors do not correct for the fact that certain observations were imputed but not observed. To correct for this, a plausible solution is to compute bootstrap estimates of the standard errors of the estimated coefficients of equation (3.3). The idea is to create 'r' replications of the original sample using as strata the industry and region. In the next step and for each replication, we estimate equation (1.4) by least squares replacing the missing observations before the new estimation of equation (3.3). From the resulting bootstrap empirical distribution of the estimators of equation (3.3), after several iterations of the replacements of the missing values, we obtain the estimated bootstrap standard errors.

b) **EM algorithm on industry, state, size variables**. Dempster, Laird and Rubin (1977) introduced the EM imputation algorithm that has been widely applied in a variety of contexts and applications. Basically, the EM algorithm imputes missing data conditional on a given model, and consequently chooses the candidate values to replace the missing cells that maximize the likelihood function conditional on the vector of parameters of that model. The purpose here is to apply all the steps of the EM algorithm to the problem at hand, using as model the standard regression model applied in the *ICA* replacement method; that is, using as covariates of the regression model the industry, region (state) and size variables. Notice that the estimation of the IC elasticities and semi-elasticities is achieved by following a iterative procedure: i) first, we apply the EM algorithm to replace the missing cells in sales, labor, capital and materials (used to compute productivity) using the procedure discussed in section 1.2 to replace missing data; ii) we compute the corresponding productivity measure (the restricted Solow residual); iii) we estimate equation (3.3) under the new imputed values of the missing observations.

c) **EM algorithm on industry, state, size variables and production function variables**. In this case we extend the set of covariates of the regression model (1.4) to include the production function variables. We follow the same iterative procedures described in case a).

d) **EM algorithm on equation (3.3).** Now we simply apply the EM algorithm to equation (3.3) and we estimate the parameters of the model in a single step by maximizing the log-likelihood function resulting from (3.3) and the MDM. The procedure now is slightly different than in cases a) and b). We now replace missing values and estimate the IC parameters in a single step by maximizing the log-likelihood function, provided the population models used to replace missing data and to estimate the IC effects on productivity are the same. Notice that the EM algorithm is always iterative; several repetitions of the replacement procedure are needed until the likelihood function converges to the conditional maximum.

We also propose mechanisms to deal with the missingness problem when we assume that the missing data mechanism (MDM) is correlated with the dependent variable of our model (non-ignorable MDM). In these cases one can implement the Heckman (1976) method (Heckit) to correct for self-selection, since OLS applied either on the complete deletion case or on the sample with replacement is inconsistent. The probability of selection—of observing the data—is modeled with the same IC of the TFP regression model plus other investment climate variables to solve the identification problem.

We also estimate equation (3.3) in the complete deletion case; that is using only the available information for PF variables without replacement (dropping all firms with missing observations from the sample). For consistency in this case we need the missing completely at random assumption (MCAR), unless we correct for the correlation between the MDM and the covariates of the model by controlling for those IC variables correlated with the MDM.

In order to compare the results of the ICA method with alternative replacement procedures we apply all these imputation models to equation (3.3). The results from the alternative imputation procedures used are reported in Table C.IV and are very robust, as we mention later on.

4. Econometric analysis of IC and productivity impact on employment, real wages, probability of exporting and probability of receiving FDI.

The same idea of approximating the unobservable fixed effect by the firm level investment climate conditions is applied in the remaining equations of the model.

The demand for labor determined by firm level productivity (*logP_{it}*) and by real wages in logs (*logW_{it}*) and is given by;

$$logL_{it} = \gamma_L + a_{L,i} + \gamma_P \log TFP_{it} + \gamma_w logW_{it} + \gamma_{Exp} y_{it}^{Exp} + \gamma_{FDI} y_{it}^{FDI} + \gamma'_{DR} D_r + \gamma'_{Ds} D_j + \gamma'_{DM} D_m + \gamma'_{DT} D_t + \varepsilon_{L,it} \quad (4.1a)$$

$$a_{L,i} = \gamma'_L IC_i^L + \gamma'_C C_i^L + v_{L,i} \; . \quad (4.1b)$$

The wage equation is determined by the productivity (TFP) level after controlling for all the IC effects and by the fact that certain firms exports and receive FDI;

$$logW_{it} = \beta_W + a_{W,i} + \beta_P \log TFP_{it} + \beta_{Exp} y_{it}^{Exp} + \beta_{FDI} y_{it}^{FDI} + \beta'_{DR} D_r + \beta'_{Ds} D_j + \beta'_{DM} D_m + \beta'_{DT} D_t + \varepsilon_{W,it} \quad (4.2a)$$

$$a_{W,i} = \beta'_{IC} IC_i^W + \beta'_C C_i^W + v_{W,i} \; . \quad (4.2b)$$

The probability of firms entering the export market depends on firm level productivity (TFP), the investment climate and by the fact that certain firms receive FDI;

$$y_{it}^{Exp} = \delta_{Exp} + a_{Exp,i} + \delta_P \log TFP_{it} + \delta_{FDI} y_{it}^{FDI} \delta'_{DR} D_r + \delta'_{Ds} D_j + \delta'_{DM} D_m + \delta'_{DT} D_t + \varepsilon_{Exp,it} \quad (4.3a)$$

$$a_{Exp,i} = \delta'_{IC} IC_i^{Exp} + \delta'_C C_i^{Exp} + v_{Exp,i} \quad (4.3b)$$

Finally, the probability of receiving foreign direct investment equation depends on firm level productivity (TFP), the investment climate and by the fact that certain firms exports;

$$y_{it}^{FDI} = \rho_{FDI} + a_{FDI,i} + \rho_P \log TFP_{it} + \rho_{Exp} y_{it}^{Exp} + \rho'_{DR} D_r + \rho'_{Ds} D_j + \rho'_{DM} D_m + \rho'_{DT} D_t + \varepsilon_{FDI,it} \quad (4.4a)$$

$$a_{FDI,i} = \rho'_{IC} IC_i^{FDI} + \rho'_C C_i^{FDI} + v_{FDI,i} \quad (4.4b)$$

Notice that since the variable y^r_{it}, with r = Exp or FDI, is a *binary random variable* taking only 0 and 1 values, then $P(y_{it}^r = 1/x) = E(y_{it}^r / x)$ then: a) the conditional probability is equal to the conditional expectation which is usually assumed to follow a Probit or a Logit model, and b) the conditional variance (heteroskedasticity) is equal to the product of the conditional probabilities of the two events. In general, the linear probability models (LPM) approximate well the Probit and Logit nonlinear models when the variables are evaluated close to their sample means. Since we are interested in the mean IC contribution relative to the mean values of the dependent variables of (4.1a) to (4.4a), we will concentrate only on linear probability specifications, like (4.3a) and (4.4a). The main advantage of the LPM is in its simplicity since the parameters of the explanatory variables of (4.3a) and (4.4a) measure the change in probability when one of the explanatory variables changes, holding the rest of the explanatory variables constant. This is important for the economic interpretation of the coefficients obtained in the empirical section.

By substituting the usually unobserved fixed effects components by their corresponding equation we can simplify the system of equations including productivity to:

$$\log TFP_{it} = \alpha_P + \alpha'_{IC} IC_i^P + \alpha'_C C_i^P + \alpha_{Exp} y_{it}^{Exp} + \alpha_{FDI} y_{it}^{FDI} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DM} D_m + \alpha'_{DT} D_t + (v_{P,i} + \varepsilon_{P,it}) \quad (4.5)$$

$$logL_{it} = \gamma_L + \gamma_P \log TFP_{it} + \gamma_w \log W_{it} + \gamma_{Exp} y_{it}^{Exp} + \gamma_{FDI} y_{it}^{FDI} + \gamma'_L IC_i^L + \gamma'_C C_i^L + \gamma'_{DR} D_r + \gamma'_{Ds} D_j + \gamma'_{DM} D_m + \gamma'_{DT} D_t + (v_{L,i} + \varepsilon_{L,it}) \quad (4.6)$$

$$logW_{it} = \beta_W + \beta_P \log TFP_{it} + \beta_{Exp} y_{it}^{Exp} + \beta_{FDI} y_{it}^{FDI} + \beta'_{IC} IC_i + \beta'_C C_i + \beta'_{DR} D_r + \beta'_{Ds} D_j + \beta'_{DM} D_m + \beta'_{DT} D_t + (v_{W,i} + \varepsilon_{W,it}) \quad (4.7)$$

$$y_{it}^{Exp} = \delta_{Exp} + \delta_P \log TFP_{it} + \delta_{FDI} y_{it}^{FDI} + \delta'_{IC} IC_i^{Exp} + \delta'_C C_i^{Exp} + \delta'_{DR} D_r + \delta'_{Ds} D_j + \delta'_{DM} D_m + \delta'_{DT} D_t + (v_{Exp,i} + \varepsilon_{Exp,it}) \quad (4.8)$$

$$y_{it}^{FDI} = \rho_{FDI} + \rho_P \log TFP_{it} + \rho_{Exp} y_{it}^{Exp} + \rho'_{IC} IC_i^{FDI} + \rho'_C C_i^{FDI} + \rho'_{DR} D_r + \rho'_{Ds} D_j + \rho'_{DM} D_m + \rho'_{DT} D_t + (v_{FDI,i} + \varepsilon_{FDI,it}) \quad (4.9)$$

The composite error terms of each equation of the system have three terms, says $\omega_{it} = \gamma_{r,i} + v_{r,i} + u_{r,it}$ with *r=P, L, W, Exp* and *FDI*. The firm fixed effects ($\gamma_{r,i}$) are approximated by the set of observed time-invariant, firm level IC and C variables. The remaining unobserved firm effects are assumed to be independently distributed of IC and C variables, therefore what remains are random effects ($v_{P,i}$). Therefore, we assume that the error terms ($v_{r,i}+\varepsilon_{r,j,it}$) are uncorrelated with all the explanatory variables of each equation *r*, where *r=P, Exp, FDI, W* and *L*. However, for certain explanatory variables this exogeneity condition is not satisfied. The endogeneity of certain IC variables induces a correlation between those IC variables and the errors ($v_{r,i}+\varepsilon_{r,j,it}$) of the system of equations (4.5) to (4.9) and creates simultaneous equation biases and inconsistencies in least squares estimators; like pooling OLS or in random effects (RE) estimators. This correlation is in general mitigated by replacing those plant-level IC variables by their region-industry averages ($\overline{IC}_j$). However, for some other explanatory variables like productivity, wages, exports and FDI, the endogeneity is intrinsic due to the simultaneous structure of the system of equations. Therefore, we estimate each equation by instrumental variables (IV) techniques (2SLS) using heteroskedasticity-robust standard errors. We could have used 3SLS, which is more efficient than 2SLS under correct specification. However, since with system of equations estimation techniques the misspecification of one equation affects the whole system, we believe that the results from 2SLS are more robust.

Provided that we are instrumenting the productivity (TFP) variable in the employment, real wages, exports and FDI equations using instruments from the investment climate survey, it is very convenient to specify a number of rules to choose the list of instruments, etc. First, estimation of the system of equations (4.5) to (4.9) by IV techniques is done equation by equation. Productivity equation is at the core of this process and it is estimated seeking robust procedures of Escribano and Guasch (2005 and 2008). Once we have obtained robust IC and C coefficients for different productivity (TFP) measures, we use the set of significant explanatory variables to instrument productivity in the rest of equations. Notice that some of these variables will be used as included instruments, while many other will be excluded instruments as they may appear as explanatory variables in other equations.

The next step is to obtain a preliminary specification for the remaining equations of the system by OLS with robust standard errors. As in the productivity case, in order to avoid omitted variables problems, the selection of the model goes from the general to the specific. We start selecting the preliminary model from a set of more than 160 IC and C variables, industry, state and year dummies, productivity and a constant term (also real wages in the case of demand for labor equation).

Once we have a preliminary valid model for each equation of the system we start instrumenting productivity. We then remove instruments from the list of excluded instruments provided we want a partial R-squared –or '*Shea*' partial R-squared—as high as possible with the restriction that our model is not over-identified. To test the over-identification restrictions we use Hansen test, a robust to general heteroskedasticity variation of classical Sargan test. In addition we take into account the significance in the first stage estimates when removing instruments. We also remove instruments from the matrix of included instruments if in the process of IV selection some of them become insignificant.

A similar process is applied when we have to instrument any other simultaneous variable like real wages in demand for labor equation, or exports or FDI when they appear as significant explanatory variable in other equations. A good strategy that works well is to estimate first by

OLS and then change to IV if we have the set of instruments, which in this case is given by the explanatory variables of the corresponding equation, excluding obviously those endogenous covariates. Then we proceed as in the productivity case, removing instruments, either included or excluded, according to the criteria mentioned before.

4.1 Identification of the system of equations

To discuss the identification issues underlying the system of equations proposed it is useful to apply matrix notation. The structural form of the system (4.5) - (4.9) is given by

$$\mathbf{A}\mathbf{y_t} + \mathbf{B}\mathbf{x_t} = \mathbf{u_t} \tag{4.10}$$

where $\mathbf{y_t}$ is the 5×1 vector of observations of *dependent* variables (log-productivity, y_{it}^{Exp} and y_{it}^{FDI}, log-employment and log-wages); $\mathbf{x_t}$ is the 140x1 vector of explanatory variables (IC_i, C_i, D_r, D_j and D_t); $\mathbf{u_t}$ is the 5×1 vector of errors; $\mathbf{A}$ is a 5×5 matrix of coefficients of simultaneous *dependent* variables; $\mathbf{B}$ is a 5x164 matrix of coefficients corresponding to the exogenous/endogenous IC and variables.

In the system (4.5) - (4.9), we are imposing certain structure; for example that employment has no direct effect in any other equation of the system and that real wages only affects employment demand, after controlling for all IC and C variables. Therefore, we can explicitly write the first LHS term of (4.10) as;

$$\mathbf{A}\mathbf{y_t} \equiv \begin{pmatrix} 1 & -a_{P,Exp} & -a_{P,FDI} & 0 & 0 \\ -a_{Exp,P} & 1 & -a_{Exp,FDI} & 0 & 0 \\ -a_{FDI,P} & -a_{FDI,Exp} & 1 & 0 & 0 \\ -a_{L,P} & -a_{L,Exp} & -a_{L,FDI} & 1 & -a_{L,W} \\ -a_{W,P} & -a_{W,Exp} & -a_{W,FDI} & 0 & 1 \end{pmatrix} \begin{pmatrix} \log TFP_{it} \\ y_{it}^{Exp} \\ y_{it}^{FDI} \\ \log L_{it} \\ \log W_{it} \end{pmatrix} \equiv \begin{pmatrix} \log TFP_{it} - a_{P,Exp} y_{it}^{Exp} - a_{P,FDI} y_{it}^{FDI} \\ y_{it}^{Exp} - a_{Exp,P} \log TFP_{it} - a_{Exp,FDI} y_{it}^{FDI} \\ y_{it}^{FDI} - a_{FDI,P} \log TFP_{it} - a_{FDI,Exp} y_{it}^{Exp} \\ \log L_{it} - a_{L,P} \log TFP_{it} - a_{L,W} \log W_{it} - a_{L,Exp} y_{it}^{Exp} - a_{L,FDI} y_{it}^{FDI} \\ \log W_{it} - a_{W,P} \log TFP_{it} - a_{W,Exp} y_{it}^{Exp} - a_{W,FDI} y_{it}^{FDI} \end{pmatrix}.$$

The rank condition is a *necessary and sufficient condition* for the system (4.10) to be identified. To discuss whether the rank condition is satisfied, say, in the first equation, let $\boldsymbol{\alpha}'$ be the first row of $\mathbf{A}$ and $\boldsymbol{\beta}'$ the first row of $\mathbf{B}$. We may now partition these vectors into two components corresponding to the included ($\boldsymbol{\alpha}_1'$ and $\boldsymbol{\beta}_1'$) variables and the excluded ($\boldsymbol{\alpha}_2'$ and $\boldsymbol{\beta}_2'$) variables in the productivity equation such that $\mathbf{A} = \begin{bmatrix} \boldsymbol{\alpha}_1' & \mathbf{0} \\ \mathbf{A_1} & \mathbf{A_2} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} \boldsymbol{\beta}_1' & \mathbf{0} \\ \mathbf{B_1} & \mathbf{B_2} \end{bmatrix}$, which allow us to construct the next matrix $\mathbf{D} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{A_2} & \mathbf{B_2} \end{bmatrix}$. By the rank condition, productivity equation is identified if $rank(\mathbf{D}) = 5-1$. The same holds for the rest of equations of the system. Thus, even if we have several exclusion restrictions in matrix $\mathbf{A}$ (in the productivity, wages and employment equations), nevertheless these restrictions are not enough to ensure that the rank condition is satisfied. For that, we force the coefficient of certain IC variables to be 0 prior to start estimating the system, for more details on extra identification issues see Escribano et al (2008b).

The empirical IC results based on 2SLS are included in Tables D.I to D.IV of the Appendix B. In all the cases we found evidence that TFP has a significant and positive impact on; employment demand, on real wages, and on the probabilities of exporting or receiving FDI. Notice that TFP is always significant even after controlling for IC and other C variables.

5. IC assessment on aggregate productivity (Olley and Pakes decomposition) and other measures of economic performance.

In the second part of the analysis, taking advantage of the robustness of the IC, and C elasticities estimated, we want to concentrate on the TFP measure that comes from the restricted Solow’s residuals. Our aim is to evaluate the IC effects on average productivity and on allocative efficiency components of the Olley and Pakes (1996) decomposition (O&P) of aggregate productivity in levels (TFP) and on the mixed O&P decomposition (logTFP).

5.1 O&P decompositions: in levels and mixed.

The O&P decomposition of aggregate productivity in levels is,

$$TFP = \overline{TFP} + N\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, TFP_{it}). \qquad (5.1a)$$

Where TFP is aggregate productivity (TFP) (or weighted average productivity, where the weights are given by the share of sales), $\overline{TFP}$ is the sample average productivity and the last term is N times the sample covariance of the share of sales and firm level productivity; this last term is the allocative efficiency term describing the ability of the markets to reallocate resources from less to more productive establishments. Furthermore, we want to exploit the log-linear properties of the following mixed[13] O&P decomposition in order to obtain closed form O&P decompositions in terms of IC and C variables,

[13] It is called mixed Olley and Pakes (O&P) decomposition because in the original O&P decomposition both TFP and the share of sales were in levels while now TFP in (5.1b) is in logs, (log P).

$$\log TFP = \overline{\log TFP} + N\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \log TFP_{it})\,. \qquad (5.1b)$$

Expressions (5.1a) and (5.1b) can be easily applied by industry, state, size, age or for the whole sample. The results of the decomposition by states and at country level in levels and in logs are in Figures 2 and 3 of appendix B.

5.2 IC effects on productivity measure in the terms of the mixed O&P decomposition.

The useful additive property of equation (3.3) in logarithms, allow us to obtain an exact closed form solution of the decomposition of aggregate log productivity according to equation (5.1b). Following Escribano et al. (2008a), we can express aggregate log productivity as a weighted sum of the average values of the IC, C, dummy D variables, the intercept and the productivity average residuals ($\bar{\hat{u}}$) from (3.3); and, the sum of the *covariances* between the share of sales and investment climate variables IC, C, dummies D and the productivity residuals ($\hat{u}$).

$$\begin{aligned}\log TFP &= \hat{\alpha}'_{IC}\overline{IC}_P + \hat{\alpha}'_C\,\overline{C}_P + \hat{\alpha}'_{DR}\,\overline{D}_r + \hat{\alpha}'_{Ds}\overline{D}_j + \hat{\alpha}'_{DM}\,\overline{D}_m + \hat{\alpha}'_{DT}\overline{D}_t + \hat{\alpha}_p + \bar{\hat{u}}_{it} \\ &+ N\hat{\alpha}'_{IC}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, IC_{P,i}) + N\hat{\alpha}'_C\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, C_{P,i}) + N_q\hat{\alpha}'_{Ds}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, D_j) + N\hat{\alpha}'_{DR}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, D_r) \\ &+ N\hat{\alpha}'_{DT}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, D_t) + N\hat{\alpha}'_{DM}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \overline{D}_m) + N\hat{\alpha}'_{DT}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \overline{D}_t) + N\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \hat{u}_{it})\end{aligned} \qquad (5.2)$$

where the set of estimated parameters used comes from the two-step TFP estimation, having the restricted Solow's residual as dependent variable of the regression equation (3.3).

The contributions of IC variables to aggregate log-TFP of equation (5.2) can be computed for the whole sample or by industry/sector, state, size, etc. In particular, we compute the IC contributions relative to aggregate productivity as follows;

$$\begin{aligned}100 = \frac{100}{\log TFP}[&\hat{\alpha}'_{IC}\overline{IC}_P + \hat{\alpha}'_C\,\overline{C}_P + \hat{\alpha}'_{DR}\,\overline{D}_r + \hat{\alpha}'_{Ds}\overline{D}_j + \hat{\alpha}'_{DM}\,\overline{D}_m + \hat{\alpha}'_{DT}\overline{D}_t + \hat{\alpha}_p + \bar{\hat{u}}_{it} \\ &+ N\hat{\alpha}'_{IC}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, IC_{P,i}) + N\hat{\alpha}'_C\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, C_{P,i}) + N_q\hat{\alpha}'_{Ds}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, D_j) + N\hat{\alpha}'_{DR}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, D_r) \\ &+ N\hat{\alpha}'_{DT}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, D_t) + N\hat{\alpha}'_{DM}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \overline{D}_m) + N\hat{\alpha}'_{DT}\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \overline{D}_t) + N\,\hat{\text{c}}\text{ov}(s_{it}^{Y}, \hat{u}_{it})].\end{aligned} \qquad (5.3)$$

There are several advantages of using equation (5.3). First, we can compare net contributions by isolating the impact of IC variables from the impact of industry dummies, the intercept, and the residuals. Second, we can split the total effect on aggregate productivity in the part explained only by IC, and C variables (demeaned *logTFP*), and the proportion is due to the rest; constant term, industry dummies and so on. The empirical results of decomposition (5.3) are in Table C.III of the appendix B.

We could also get rid of the different directional effects (positive or negative) of the various IC effects by simply computing the percentage contributions in absolute value. This slightly modification allow us to do direct comparisons of the IC absolute percentage contributions (or weight of each IC variable relative to the total weight of other IC variables) to aggregate log-productivity, to average log-productivity and to the allocative efficiency term. The results are in Figure 4.1.

5.3 Simulations based on the IC effects on the O&P decomposition of TFP.

So far, we have exploited the linear properties of the logarithm form of the mixed O&P decomposition of TFP. However, the original O&P decomposition was done in terms of TFP and the share of sales (in levels). Therefore the O&P decomposition is capturing also nonlinear relations between market shares and IC variables coming from (5.1a) and equation (3.3). To know to what extent these nonlinear terms are affecting this relation, we perform simulation experiments[14] on INF, IC, and C variables, and evaluate the consistency of the results with the ones obtained from the previous mixed O&P decomposition- see (5.3). The IC simulations are done variable by variable (one at a time) keeping the rest of the variables constant; that is, we propose a scenario in which one of the IC variables experiment a 20 percent improvement in all the establishments. We compute the corresponding rate of change of aggregate productivity, average productivity and allocative efficiency caused by such improvement. We repeat the same experiment for the rest IC and C variables, and, for comparative purposes, we also evaluate the relative IC effect by group of IC variables.

The resulting simulations of a 20% improvement in IC variables are in Figure 4.2. A comparison between the simulations and the IC absolute percentage contributions are in Figure 4.3.

5.4 International comparisons of IC effects on aggregate demeaned productivity.

To make cross-country comparisons based on IC impacts on productivity, avoiding the problem of comparing apples and oranges, it is desirable to create an index (*demeaned productivity*). After subtracting the mean (that is, the constant term, time effects, industry effects

[14] We are indebted to Ariel Pakes for this suggestion.

and country-specific effects) from firm level log-productivity we can concentrate on the part of log-productivity explained by the IC variables. Thus, demeaned aggregate log-productivity at the firm level is simply

$$Demeaned \log TFP_i = \hat{\alpha}'_{IC} IC_i^P + \hat{\alpha}'_C C_i^P \quad (5.4)$$

With the expression given by (5.4) we can easily compute the O&P decompositions (5.1a) and (5.1b) based on this demeaned part of productivity to do international comparisons of IC impacts on aggregate productivity. The share of aggregate log-productivity attributable to the IC can be found at the end of Table C.III, while a comparison of the demeaned productivity in India with that of other countries is in Figure 9.

5.5 IC evaluation on the sample means of employment and wages, on the probability of exporting and on the probability of receiving FDI.

The objective now is to measure the *partial direct effect* of each IC variable on each dependent measuring economic performance from the system of equations (4.5)-(4.9), at different aggregation levels (aggregate level, by sector, by region, by size of the firm, by age of the firm, etc.). For that purpose, we evaluate the impact of the average IC variable on the sample average values of the dependent variables of the system. In what follows, we substitute all the unknown parameters of the system (4.5) to (4.9) by their corresponding 2SLS estimated values.

The labor demand and the wage equations evaluated at the sample means and in relative terms are,

$$\begin{aligned} 100 = \frac{100}{\overline{\log L_t}} [& \hat{\gamma}_L + \hat{\gamma}_P \overline{\log TFP}_t + \hat{\gamma}_w \overline{\log W}_t + \hat{\gamma}_{Exp} \overline{y}_t^{Exp} + \hat{\gamma}_{FDI} y_t^{FDI} + \hat{\gamma}'_L \overline{IC}^L \\ & + \hat{\gamma}'_C \overline{C}^L + \hat{\gamma}'_{DR} \overline{D}_r + \hat{\gamma}'_{Ds} \overline{D}_j + \hat{\gamma}'_{DM} \overline{D}_m] \end{aligned} \quad (5.5)$$

$$\begin{aligned} 100 = \frac{100}{\overline{\log W}_t} & \hat{\beta}_W + \hat{\beta}_P \overline{\log TFP}_t + \hat{\beta}_{Exp} y_t^{Exp} + \hat{\beta}_{FDI} y_t^{FDI} + \hat{\beta}'_{IC} \overline{IC}^W \\ & + \hat{\beta}'_C \overline{C}^W + \hat{\beta}'_{Ds} \overline{D}_j + \hat{\beta}'_{Ds} \overline{D}_j + \hat{\beta}'_{DR} \overline{D}_r + \hat{\beta}'_{DM} \overline{D}_m \end{aligned} . \quad (5.6)$$

Since y_{it}^{Exp} and y_{it}^{FDI} are binary variables, evaluating the impact at the sample mean implies the evaluation on the *probability (frequency) of exporting and receiving FDI*, respectively. In particular equations (3.8) and (3.9) relative to the frequency of exporting and receiving FDI becomes

$$100 = \frac{100}{\hat{P}(Exp_t \succ 0)} \hat{\delta}_{Exp} + \hat{\delta}_P \overline{\log TFP}_t + \hat{\delta}_{FDI} y_{it}^{FDI} + \hat{\delta}'_{\mathrm{IC}} \overline{\mathrm{IC}}^{Esp} + \hat{\delta}'_C \bar{C}^{Exp} + \hat{\delta}'_{Ds} \bar{D}_j + \hat{\delta}'_{DR} \bar{D}_r + \hat{\delta}'_{DM} \bar{D}_m \quad (5.7)$$

$$100 = \frac{100}{\hat{P}(FDI_t \succ 0)} \left[\hat{\rho}_{FDI} + \hat{\rho}_P \overline{\log TFP}_t + \rho_{Exp} \bar{y}_t^{Exp} + \hat{\rho}'_{\mathrm{IC}} \overline{\mathrm{IC}}^{FDI} + \hat{\rho}'_{\mathrm{C}} \bar{C}^{FDI} + \hat{\rho}'_{\mathrm{Ds}} \bar{D}_j + \hat{\rho}'_{DR} \bar{D}_r + \hat{\rho}'_{DM} \bar{D}_m \right] \quad (5.8)$$

The results of equations (5.5) to (5.8) are in Figures 5.1 to 5.4 of appendix B.

6. Empirical results

6.1 Key Results on Productivity

Figure 6 in its first column compare the relative importance of groups of IC variables in terms of contributions to average log-productivity at the aggregate level. Red tape, informality and others factors represent 32.1% of the whole contribution of IC and C variables to average log-productivity; within these factors, productivity is affected (see Figure 4.1) by bureaucratic constraints such as having interventionist labor regulation; informalities in the relations between firms and the government such as having the advantage of speed up bureaucracy via informal payments or the possibility of maintain undeclared a percentage of sales or workforce to tax authorities; in the other hand, if firms have security costs or related with absenteeism of their employees their productivity could be affected. The largest contributions to average log-productivity within this group come from the percentage of workforce declared to taxes and the dummy for security, as Figure 4.1 shows.

Other control variables group is the second in order of importance representing 22.1% of the whole contribution of IC variables to average log-productivity (see Figure 6). Figure 4.1 breaks down this percentage in key factors: being an incorporated company, the age of the firm, the share of direct exports, the percentage of workforce unionized, the number of days losses due to strikes and the dummy for medium firms. The largest contributions are given by the age of the firm.

Next group in order of importance is finance and corporate governance, with a relative impact on average log-productivity of 21.4%, as Figure 6 shows. Four variables within this group affect on productivity equation: belonging to a commerce chamber, percentage of firm's working capital financed with funds from domestic commercial banks, if the firm has access to a loan line and if the firm has its annual statements externally audited, having dummy for external audit the highest contribution. (See Figure 4.1)

Infrastructures relative importance with respect to all IC variables in productivity equation is 14.2%, as Figure 6 shows. Within this group firms' productivity is affected by the longest time spent dealing with customs to export, having an own generator, the percentage of water supply from public source, the value of shipment losses in the domestic market, having own transport and if firms use regularly web page to communicate with their clients and suppliers . The largest contributions come from dummy for own generator and water from public sources variables.

The last group in order of importance is quality, innovation and labor skills which relative weight within all IC variables is only of 10.2% see Figure 6. The factors of this group affecting productivity are: if the plant has introduced any product innovation, if the firm has performed R&D activities, if the plant uses technology licensed from a foreign company, the percentage of workforce receiving internal training, the percentage of unskilled workers and the percentage of staff using computer at job. The largest contributions to average log-productivity within this group come from workforce with computer and dummy for R&D. (see Figure 4.1).

Figure 4.4 stress the differences of the impact of IC groups of variables on productivity by size classification. The most relevant issues are the greater contribution of red tape, informality and others for the case of small firms and the more important contribution of quality, innovation and labor skills in large firms.

We focus now on the decomposition of the allocative efficiency term (or covariance term) of the Olley and Pakes decomposition in logs. Column contributions of the efficiency term section of Figure 4.3 shows the relative impact of each group of IC and C variables on this term at the aggregate level. The main group affecting the allocative efficiency is other control variables representing 26.6% of the whole contribution of IC and C variables to the allocative efficiency. Next group is quality innovation and labor skills being its weight 24.5%, followed by infrastructures which weight is 21.4%. The relative contributions of finance and corporate governance and red tape, informality are 19.3% and 8.2% respectively. Figure 4.1 shows what variables have the largest contributions on the allocative efficiency; the most important contribution comes from the dummy for foreign license and the dummy for own generator and water from public sources variable has a high impact too.

6.2 Key Results on Employment

Figure 6 in its second column compares the relative importance of groups of IC and C variables in terms of contributions to average log-employment at the aggregate level. The largest relative impact comes from real wages. Its contribution is 44.5%, moreover its effect is negative as Table shows, obviously the demand of employment decreases as wages increases. It must be pointed out that the percentage contribution of real wages in Figure 5.1 is the largest among all IC and C variables.

Quality, innovation and labor skills weight is 18.7% (see Figure 6). Six variables have impact within this group: having a quality certification, if the firm has performed R&D activities, the percentage of workforce receiving internal training, the percentage of female workers, the percentage of skilled workers and the percentage of staff using computer at job. Figure 5.1 shows that staff – skilled workers has the largest contribution on employment demand.

Other control variables, representing 11.3% of the whole contribution of IC and C groups to average log-employment. The factors of this group involving the demand of employment are presented in Figure 5.1 and they are commented in what follows: the age of the firm, being an incorporated company, share of exports and the percentage of workforce unionized. The largest contributions of this group of variables come from the age of the firm and from being an incorporated company.

The next group in order of importance is finance and corporate governance, with a relative impact on average log-employment of 8%, as Figure 6 shows. The percentage of firm's capital owned by the largest shareholder, belonging to a commerce chamber, having an overdraft facility and the percentage of total borrowing denominated in foreign currency have impact on average log-employment. The largest contribution within this group comes from the percentage of firm's capital owned by the largest shareholder, as Figure 5.1 illustrates.

Productivity relative weight is 6.8% the third lowest among groups of variables, nevertheless when its contribution to average log-employment is compared to other variables separately its relative importance grows (see figure 5.1).

In what refers to red tape, informality and others group its relative weight in Figure 6 is 6.3%. Figure 5.1 describes the factors of this group affecting employment: the percentage of sales reported to taxes and having security costs. The largest contribution of this group is given by the percentage of sales reported to taxes.

Regarding infrastructures group its relative weight in Figure 6 is only 4.5%. The infrastructure factors affecting the demand of employment are (see Figure 5.1) having own generator, the shipment losses in domestic market, having own transport and if firms use regularly web page to communicate with their clients and suppliers. Finally, the largest contribution comes from the dummy for own generator.

Figure 7.1 illustrates the differences among sizes in the decomposition of average log-employment.

6.3 Key Results on Real Wages

Third column of Figure 6 illustrates the relative importance explaining average log-wages of each group of IC and C variables. Quality, innovation and labor skills group represents 27.1% of the whole contribution of IC and C variables to average log-wage. Specifically, the factors of this group that has an effect on wages are shown in Figure 5.2, these factors are: if firm has performed R&D activity, the percentage of skilled workforce, having an external training program and the years of tenure of the manager; the largest contributions come from dummy for external training and staff - skilled workers variables.

The second most important contribution comes from log-productivity with an impact of 20.4%. Real wages are closely and positively related with productivity. Productivity relative importance becomes even larger when compared individually with other IC factors as Figure 5.2 shows.

Finance and corporate governance group is behind productivity in order of importance, being its relative weight in Figure 6 17.5%. Figure 5.2 list the factors of this group; the percentage of firm's capital owned by the largest shareholder, having a loan and the percentage of fixed assets financed by domestic commercial banks.

Next group is other control variables; the relative weight of this group in Figure 6 is 15.2%. Figure 5.2 enumerates the specific factors of this group: being an incorporated company, the number of competitors and the share of direct imports. Figure 5.2 also illustrates the relative importance of each factor in terms of percentage contributions to average log-wage; the number of competitors has the largest impact.

The relative weight of the red tape, corruption and crime group in Figure 6 is 14.9%. Figure 5.2 highlights which concrete factors of this group have effect on wages. The percentage of sales reported to taxes, suffering criminal acts, realizing illegals payments for protection and manager´s time spent in bureaucratic issues have impact on wage, having the percentage of sales reported to taxes the largest contribution.

Within infrastructures three factors are related with wages, with a relative importance smaller than other groups, as Figure 6 shows. Having own generator, having own transport and the value of shipment losses in the domestic market. The largest contribution comes from dummy for own generator.

Figure 7.2 explores the differences among the decomposition of average log-wage by sizes.

6.4 Key Results on the Probability of Exporting

Red tape, informality and others group is of key importance for exports as Figure 6 in its fourth column shows; this group explains 31.3% of the whole impact of IC and C variables on the probability of exporting. Those firms expending money in security have more probability of being exporter. Bureaucratic constraints and informalities limit exporting activities as well, as impact of having an interventionist labor regulation and realizing payments to speed up bureaucracy. The largest impact is given by dummy for interventionist labor regulation variable, as Figure 5.3 shows.

Productivity is behind red tape, informality and other factors in order of importance as Figure 6 shows, being its relative weight 26.8%. The marginal effect of productivity on the probability of exporting is positive and large, the more productivity the more probability of becoming exporter. From Figure 5.3 it is clear that when compared individually with other IC factors productivity become even more important, being only exceeded by one IC variables in terms of percentage contributions to the probability of exporting.

Other control variables group weight in Figure 6 is 17%. Within this group, two variables have a significant impact on the probability of exporting: the age of the firm and the number of competitors. Both variables have an important contribution.

Next group in order of importance is finance and corporate governance group, its relative weight in Figure 6 is 14.1%. Three variables within this group affect on export equation: the percentage of firm’s working capital financed with funds from domestic commercial banks, having a current account and the percentage of total borrowing denominated in foreign currency. From Figure 5.3 we are able to rank finance factors by their percentage contribution to the probability of exporting, the main contribution comes from the dummy for current account variable.

The relative contribution in Figure 6 of infrastructures group is 8.7%. Improved infrastructures make easier to export; if firms are dealing with large waiting times in customs their probability of exporting becomes smaller. Having own generator, having own transport and using of IC technologies such a web page or e-mail when doing business, also have impact on probability of exporting. From Figure 5.3 we are able to identify which are the factors with the largest impact on the probability of exports.

The lowest contribution comes from quality, innovation and labor skills group; the relative weight of this group in Figure 6 is 2.1%. Figure 5.3 highlights which concrete factors of this group have effect on probability of exporting. If the firm has performed R&D activities, if the plant has performed outsourcing and the percentage of workforce receiving external training..

The contributions of IC and C variables on the probability of exporting among sizes are illustrated in Figure 7.3.

6.5 Key Results on the Probability of Receiving Foreign Direct Investment

We now focus on the results of the foreign direct investment equation. Last column of Figure 6 shows that productivity is a key factor affecting FDI decisions; its weight in Figure 6 is 26.6%. Its effect is positive, meaning that more productivity implies more probability of receive FDI. From Figure 5.4 it is clear that productivity has the largest contribution to the probability of receiving FDI among all IC and C variables.

Regarding quality, innovation and labor skills its relative weight in Figure 6 is 26.8%, the most important impact among IC groups. Specifically, the factors of this group that has an effect on the probability of receiving FDI are shown in Figure 5.4, these factors are: if the plant uses technology licensed from a foreign company, the percentage of unskilled workers, the percentage of workforce receiving external training, the number of years of experience of manager and the percentage of workers who use computer; the largest contribution comes from the experience of the manager.

Finance and corporate governance is the next group in order of importance, with a relative weight in Figure 6 of 18.4%. The finance factors that affect positively the probability of receiving FDI are: belonging to a commercial chamber, the percentage of firm’s working capital financed by domestic private banks and having an overdraft facility. The largest contribution of this group is given by the percentage of firm’s working capital financed by domestic private banks (see Figure 5.4).

Other control variables group has a weight in Figure 6 of 11.7%. The age of the firms and number of days of production losses due to strikes are the only two factors with effect on the probability of receiving FDI, as Figure 5.4 shows.

Infrastructures group is the next group in order of importance in Figure 6, being its relative weight 10.7%. Within this group, the number of power outages, having own generator and having own transport have impact on probability of receiving FDI. The main contribution of this group in Figure 5.4 is given by the number of power outages.

Red tape, informality and others factors are the last in order of importance explaining the probability of receiving FDI; their joint relative weight in Figure 6 is only of 5.8%. Three variables within this group affect on FDI equation: dummy for conflicts with court involved, dummy for criminal activity and dummy for illegal payments for protection. The contributions of three variables are similar and low, as Figure 5.4 shows.

The sensitivity to IC factors of the probability of receiving FDI does not vary dramatically among sizes, see Figure 7.4.

6.7 Key Results on the Manager's Perceptions

Figures 8.1 and 8.2 shown that the three most important obstacles in order of relevance are: Electricity, Taxes (high rates or tax administration) and corruption.

Figure 8.1 in its sixth column shows the average group relative weights of four IC variables groups: infrastructures, red tape, informality and others, finance and labor skills. Having red tape, informality and others group the largest relative weight. Infrastructures group has an important relative weight too.

6.8 International Comparison.

We compare India's performance with ten selected economies: Brazil, Chile, Mexico, Turkey, Philippines, Indonesia, Egypt, Bangladesh, Croatia and Pakistan.

Additional insight on the role of the investment climate on productivity is obtained by applying the O&P decomposition on the concept of *demeaned* productivity; the share of productivity associated only to investment climate variables. Figure 9.1 compares the O&P *demeaned* decomposition of India with those of other countries.

Figure 9.1 says that India aggregate log-productivity is in general positively influenced by the IC. This does not mean that India is more productive than other countries, but that the effect of the IC on productivity is larger in India and that the positive IC factors dominate over the negative IC ones. The dominant contributor is the average productivity.

7. Conclusions

Large differences in output per worker between rich and poor countries are usually attributed to differences in Total Factor Productivity (TFP). Thus, in order to get the objectives of increasing productivity and reduce unemployment, the main objective of Indian's economic authorities is seeking ways to stimulate country competitiveness and TFP is usually its main driver.

In this paper we have extended the robust productivity (TFP) approach of Escribano and Guasch (2005, 2008) to alternative Olley and Pakes (1996) decompositions of TFP, logTFP and the mixed case. Each of them has certain advantages and disadvantage over the others. First, the O&P decomposition of TFP has the advantage that is the natural measure of productivity, measures linear and nonlinear relationships between TFP and IC variables, but has the drawback that it is difficult to get closed form relationships between TFP and IC and we have to obtain it by simulation methods. Second, the O&P decomposition in logs provides exact decompositions relating IC with logs TFP in terms of average productivity as well as with the allocative efficiency term. It also allows us to obtain an explicit decomposition relating the inputs (L, M and K) with the IC variables. It has the disadvantage, that they only provide approximate results since we are not interested in the IC effects on log TFP but on the effects on TFP. Finally, the mixed decompositions provide a nice compromise between the two. It gives us an explicit relationship between IC and Log TFP but also one in terms of efficiency. Furthermore, the allocative efficiency term is measured in terms of the share of sales and not the share of log sales, which is not invariant to changes on the units of measurements on sales. The empirical results in terms of the O&P are very similar to those in terms of the mixed TFP. However, with the O&P in logs the allocative efficiency effect on aggregate productivity is much lower in absolute terms.

We have proposed to use a *demeaned O&P decomposition* for the evaluation of the IC effects by blocks and proposed to compare those proportions with the ones obtained from firm's perception on bottlenecks for economic performance of firms. In fact, we obtain empirical results that are consistent with firm's perceptions. The most important block of IC variable in Indian is red tape, informality and others with the main individual IC constraints being; are taxes and tax administration and security. The second and third groups of IC variables are other control variables and infrastructures with similar contributions. The most important single elements within these blocks are: share of exports, having own generator and the number of days to clear custom to export. Those results are also consistent with the ones obtained in term of ease of doing business of DBR (2007).

This new measure of TFP allows us to make interesting cross country comparisons. We found that India has a middle level of demeaned productivity among the countries considered.. The corresponding ranking of countries in term of demeaned O&P decompositions are highly correlated with per-capita income rankings and with firm perceptions or the rankings base on the ease of doing business.

We believe that improving the investment climate (IC) is a key policy instrument to promote economic growth and to mitigate the institutional, legal, economic and social factors that are constraining the convergence of per capita income and labor productivity of India relative to more developed countries.

In this paper, we also identify the main investment climate variables that affect economic performance measures like total factor productivity, employment, wages, exports and foreign direct investment. We extend the productivity methodology of Escribano and Guasch (2005, 2008) and Escribano et al. (2008), based on the analysis of how the investment climate affect productivity, to other economic performance measures. We have proposed a system of five simultaneous equations to analyze the interactions between TFP and other economic performance measures.

We found that TFP is a key variable explaining other important economic decisions for the firm, like employment demand, wages, exports and FDI, even after controlling for the investment climate environment.

From the analysis of Firm's perceptions, we identify the block of red tape, informality and others as the main IC block creating severe obstacles for firm economic performance. The main individual IC bottlenecks within this group are taxes and tax administration. The second is infrastructures group being electricity the most serious obstacle among all IC bottlenecks. The third and fourth IC blocks are the block of quality innovation and labor skills and finance.

The Doing Business report (2007), DBR, identifies three main problems; enforcing contracts, paying taxes and trading across borders.

From our econometric analysis we observe similar results since red tape, corruptions and crime is the main issue in terms of productivity and exports. The main IC variables from this group are: dummy for interventionist labor regulation and taxes. This is also consistent with the DBR. The main econometric effect on productivity from the IC block on infrastructures is also the longest number of days to clear customs for exports.

TFP in Indian's manufacturing firms is very important to enhance international trade. Not only it affects the capacity of firms to export but also affects the probability of the firms to attract foreign direct investment (FDI), having the largest contribution of all IC groups. The quality of infrastructure in India also affects the probability of receiving FDI with the number of power outages being an important individual determinant.

We conclude that for policy analysis it is very useful to combine different sources of information; firm perceptions on bottlenecks, ease of doing business conclusions form DBR and the econometric performance analysis based in investment climate surveys.

References

Ackerberg D. A. and K. Caves (2003). “Structural Identification of Production Functions”. *Working Paper,* University of California, Los Angeles.

Alexander, W. Robert J., Bell, John D., and Knowles, Stephen (2004). “Quantifying Compliance Costs of Small Businesses in New Zealand”. Discussion paper, University of Otahgo.

Barro R. J. and X. Sala-i-Martin (2004). Economic Growth. The MIT Press, second edition, Cambridge, Massachusetts.

Bartelsman E. J. and M. Doms (2000). “Understanding Productivity: Lessons from Longitudinal Microdata”. *Journal of Economic Literature*, Vol. 38, Nº 3, 569-594.

Bosworth, Barry and Susan Collins (2003). “The Empirics of Growth: An Update”. The Brookings Institution. Washington, D.C. Processed.

Cole, H. L., L. E. Ohanian, A. Riascos and J. A. Schmitz Jr. (2004). "Latin America in the Rearview Mirror". National Bureau of Economic Research WP #11008, December.

Diewert W. Erwin and Alice O. Nakamura (2002) “The Measurement of Aggregate Total Factor Productivity Growth”. J. J. Heckman and E. E. Leamer (eds.). Handbook of Econometrics, Vol. 6, forthcoming.

Djankov, S. and Hoekman, B., (2000). Foreign Investment and Productivity Growth in Czech Enterprises, *World Bank Economic Review*, 14, 49-64

Djankov, Simeon, La Porta, Rafael, Lopez-de-Silanis, Florencio, and Shleifer, Andrei (2002). “The Regulation of Entry”. *Quarterly Journal of Economics* 117, February 2002. 1-37.

Escribano Alvaro and J. Luis Guasch (2005); “Assessing the Impact of the Investment Climate on Productivity using Firm Level Data: Methodology and the Cases of Guatemala, Honduras and Nicaragua,” *World Bank Policy Research Working Paper #* 3621, The World Bank, June.

Escribano A. and J. L. Guasch (2008); “Robust Methodology for Investment Climate Assessment on Productivity: Application to Investment Climate Surveys from Central America,” *Working Paper # 08-19 (11)*, *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena. (2008a); “Investment Climate and Firm’s Performance: Econometric and Applications to Turkey’s Investment Climate Survey,” Working Paper # 08-20 (12), *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena (2008b); “Investment Climate Assessment Based on Demeaned Olley and Pakes Decompositions: Methodology and Applications to Turkey’s Investment Climate Survey,” Working Paper # 08-20 (13), *Universidad Carlos III de Madrid*.

Escribano A., N. Peltier, L. Garrido and H. Sing (2005b). “The Impact of Infrastructure on Competitiveness in Latin America: A Firm Level Analysis Based on Investment Climate

Assessments". *Paper presented at the Joint Conference of the Interamerican Development Bank (IADB) and The World Bank (WB)*, June 6-7.

Escribano, A. and J. Pena (2008); "Empirical Econometric Evaluation of Alternative Methods of Dealing with Missing Values in Investment Climate Surveys," mimeo, *Universidad Carlos III de Madrid*.

Escribano, A., and R. Stucchi (2008). "Catching Up in Total Factor Productivity through the Business Cycle: Evidence from Spainish Manufacturing Surveys," Working Paper #08-51, Universidad Carlos III de Madrid.

Foster L., J. Haltiwanger and C.J. Krizan (1998). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence". *NBER Working Paper* W6803.

Guasch, J. Luis. (2004). "An Assesment of Logistic Costs and of their Impact on Competitiveness", World Bank, Washington, DC.

Hall R. E. (1990). "Invariance Properties of Solow´s Productivity Residual". In Peter Diamond (ed.). Growth, Productivity, Employment. Cambridge: MIT Press, 1-53.

Haltiwanger, John (2002). "Understanding Economic Growth. The Need for Micro Evidence". *New Zealand Economic Papers* 36 (1), 33-58.

He, Kathy S., Morck, Randall and Yeung, Bernard (2003). "Corporate Stability and Economic Growth". William Davidson Working Paper No. 553.

Hulten Ch. R. (2001). "Total Factor Productivity: A Short Biography". In Ch. R. Hulten, E.R. Dean and M. J. Harper (eds.) New Developments in Productivity Analysis, The University of Chicago Press, 1-47.

Jorgenson D.W. (2001). "Information Technology and the U.S. Economy". *The American Economic Review*, Vol. 91, 1, 1-32.

Jorgenson D.W., F. Gollop and B. Fraumeni (1987). Productivity and U.S. Economic Growth. Cambridge: Harvard University Press.

Kasper, Wolfgang, (2002), "Losing Sight of the Lodestar of Economic Freedom: A Report Card on New Zeland's Economic Reform", NZ Business Roundtable.

Kerr, Roger (2002). "The Quagmire of Regulation" NZ Business Roundtable.

Levinsohn J. and A. Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables". *Review of Economic Studies*, 70, 317-341.

Loayza, N.V., A. M. Oviedo and Luis Serven (2004). "Regulation and Macroeconomic Performance", World Bank, Washington DC.

Mitra D. and B. P. Ural (2007), "Indian Manufacturing: a Slow Sector in a Rapidly Growing Economy". World Bank, Washington DC.

McMillan, John, (1998), " Managing Economic Change: Lessons from New Zeland", *The World Economy 21,* 827-43.

McMillan, John (2004), “A Flexible Economy? Entrepeneurship and Productivity in New Zeland”, Working Paper, Graduate Scholl of Business, Stanford University, Stanford, CA.

Olley G. S.and A. Pakes (1996). “The Dynamics of Productivity in the Telecommunications Equipment Industry”. *Econometrica*, Vol. 64, 6, 1263-1297.

Organization for Economic Cooperation and Development (2001). *Businesses’ Views on Red Tape,* Paris, OECD.

Prescot, Edward C. (1998) "Needed: A Theory of Total Factor Productivity". International Economic Review, 39, 525-552.

Rodrik, Dani, and Arvind Subramanian (2004). “From ‘Hindu Growth to Productivity Surge: The Mystery of the Indian Growth Transition”. Harvard University, Cambridge, Mass. Processed.

Solow R. M. (1957). “Technical Change and the Aggregate Production Function”. The *Review of Economics and Statistics*, 39 (3), 312-320.

Veeramani, C., and G. Goldar. (2004). “Investment Climate and Total Factor Productivity in Manufacturing: Analysis of Indian States.” Working paper #127, Indian Council for Research on International Economic Relations, New Delhi, India.

World Bank (2005): “World Development Report 2005: A Better Investment Climate for Everyone,” *The World Bank and Oxford University Press*, Washington, D.C.

World Bank (2009): “Doing Business Report 2009: Comparing Regulation in 181 Economies,” *The World Bank Group*, Washington DC, USA.

World Bank 2007. “Unleashing India`s Innovation: Toward Sustainable and Inclusive Growth”. Washington, D.C. World Bank

World Bank 2009. “India`s Investment Climate: Voices of Indian Business”. Washington, D.C. World Bank

World Economic Forum, The (2008).”Global Competitiveness Report”

Wilkinson, Bryce (2001). “Constraining Government Regulation”. NZ Business Roundtable.

Appendix A: Definitions

I Production Function Variables[15]

Sales: Used as the measure of output for the production function estimation. Sales are defined as total annual sales. The series are deflated by using the Producer Price Indexes (PPI), base 2002.

Employment: Total number of permanent and temporal workers.

Total hours worked per year: Total number of employees multiplied by the average hours worked per year.

Capital stock: Total costs of intermediate and raw materials used in production (excluding fuel). The series are deflated by using the Producer Price Indexes (PPI), base 2002.

User cost of capital: The user cost of capital is defined in terms of the opportunity cost of using capital; it is defined as a 15% of the net book value of machinery and equipment.

Labor cost: Total expenditures on personnel. The series are deflated by using the Producer Price Indexes (PPI), base 2002.

II Dependent Variables in Equation Regressions and Linear Probability Models

Demand for Labor: Total number of permanent and temporal workers.

Real Wage: Real wage is defined as the total expenditures on personnel (deflated by using the Producer Price Indexes (PPI), base 2002.) divided by the total number of permanent and temporal workers.

Export: Dummy variable that takes value 1 if exports are greater than 10%.

Foreign Direct Investment: Dummy variable that takes value 1 if any part of the capital of the firm is foreign.

III General Information at Plant Level

Industrial classification: a) food; b) apparels; c) textiles and leather; c) chemicals and chemical products; d) plastics and rubbers; e) non-metallic products; f) structural metal and metal products; g) machinery and equipment.

Regional classification: a) Andhra Pradesh; b) Bihar; c) Delhi; d) Gurajat; e) Haryana; f) Jharkhand; g) Karnataka; h) Kerala; i) Madhya Pradesh; j) Maharashtra; k) Orissa; l) Punjab;
m) Rajasthan; n) Tamil Nadu; o) Uttar Pradesh; p) West Bengal.

Size classification: a) small firms (< 20 employees); b) medium firms (>=20 & <100); c) large firms (>=100).

[15] All series figure in US dollars, data obtained from WDI, The World Bank, 2008.

Table A.I: Definitions of Investment climate (IC) variables of Infrastructures group.

Name of the variable	Description of the variable	Observations (Response rate)
Days to clear customs to export	Average number of days to clear customs when exporting directly.	1275 (18.6)
Longest days to clear customs to export	Longest number of days to clear customs when exporting directly.	1260 (18.4)
Days to clear customs to import	Average number of days to clear customs when importing.	453 (6.6)
Longest days to clear customs to import	Longest number of days to clear customs when importing.	441 (6.4)
Power outages	Total number of power outages suffered by the plant in 2004.	6525 (95.3)
Average duration of power outages	Average duration of power outages suffered in hours, conditional on the pant reports having power outages.	5049 (73.7)
Losses due to power outages	Losses due to power outages as a percentage of total annual sales, conditional on the plant reports having power outages.	5016 (73.2)
Wait for electricity supply	Number of days waiting to obtain an electricity supply, conditional on submit an electrical connection.	2301 (33.6)
Dummy for gifts electric supply.	Gifts expected or requested to obtain an electrical connection, conditional on submit an electrical connection.	2181 (31.8)
Dummy for own generator	Dummy variable taking value 1 if the firm has its own power generator.	6648 (97.1)
Electricity from a generator	Percentage of the electricity used by the plat provided by an own generator.	3372 (49.2)
Water outages	Total number of water outages suffered by the plant in 2004.	5784 (84.5)
Water supply from public sources	Percentage of water supply from public sources.	6504 (95)
Wait for water supply	Number of days waiting for a water supply, conditional on submit a water supply.	1647 (24)
Dummy for gifts for water supply	Gifts expected or requested to obtain a water supply, conditional on submit a water supply.	1464 (21.4)
Wait for phone connection	Number o days waiting to obtain a phone connection, conditional on submit a phone connection.	2541 (37.1)
Dummy for gifts for phone connection	Gifts expected or requested to obtain a phone supply, conditional on submit a phone connection	2424 (35.4)
Dummy for web page	Dummy variable taking value 1 if the plant uses its own web page to communicate with clients and suppliers.	6819 (99.6)
Dummy for e-mail	Dummy variable taking value 1 if the plant uses the electronic mail to communicate with clients and suppliers.	6816 (99.5)
Dummy own transport	Dummy variable taking value 1 if the plant uses its own transport to make shipments to its customers.	6786 (99.1)
Shipment losses in domestic market	Percentage of the consignment value of the products shipped lost while in transit because of theft, breakage or spoilage in the domestic market	6396 (93.4)
Shipment losses in international market	Percentage of the consignment value of the products shipped lost while in transit because of theft, breakage or spoilage in the international market.	4605 (67.2)

Table A.II: Definitions of Investment climate (IC) variables of Red tape, informality and others group.

Name of the variable	Description of the variable	Observations (Response rate %)
Sales reported to taxes	Percentage of total annual sales that a typical firm operating in plant's sector reports for tax purposes.	6591 (96.2)
Workforce reported to taxes	Percentage of total work force that a typical firm operating in plant's sector reports for tax purposes.	6534 (95.4)
Dummy for conflicts in courts	Dummy taking value 1 if the plant has conflicts with clients with a court involved (conditional on having conflicts with clients with a third part involved).	6507 (95)
Dummy for security	Dummy taking value 1 if the plant has security expenses.	6525 (95.3)
Dummy for crime	Dummy taking value 1 if the plant has experienced losses due to criminal attempts in 2004.	6723 (98.2)
Dummy for illegal payments for protection	Dummy taking value 1 if the plant has experienced cost due to protection payments, e. g. to organized crime, to prevent violence (bribery).	6312 (92.2)
Manager's time in bureaucratic issues	In typical week percentage of manager's time spent dealing with bureaucratic issues.	6795 (99.2)
Number of inspections	Total number of inspections of tax officials received by the plant in 2004.	5667 (82.7)
Average duration of inspections	Average duration of inspections in hours.	4950 (72.3)
Dummy for gifts in inspections	Gifts expected or requested in inspections with tax officials.	1914 (27.9)
Dummy payments for contract with the government	Dummy that takes value 1 if firms operating in the same sector of the surveyed plant have to offer informal payments to obtain a contract with the government.	6726 (98.2)
Payments to obtain a contract with the government	Payments to obtain a contract with the government as a percentage of contract value.	1800 (26.3)
Dummy for payments to speed up bureaucracy	Gifts or informal payments to public officials to "get things done" with regard to customs, taxes, licenses, regulations, services etc.	6795 (99.2)
Dummy for interventionist labor regulation	Dummy variable that takes value 1 if the labor regulation has affected plant's employment decisions.	6726 (98.2)
Overdue payments	Number of days to resolve overdue payments.	3915 (57.2)
Sales never repaid	Percentage of monthly total sales to private customers that were never repaid.	3942 (57.6)
Wait for a construction permit	Days waiting to obtain a construction permit (conditional on submit a construction permit).	1680 (24.5)
Dummy for gifts to obtain a construction permit	Gifts expected or requested to obtain a construction permit, conditional on submit a construction permit.	1503 (21.9)
Wait for an operating license	Days waiting to obtain a main operating license (conditional on submit a operating license).	1659 (24.2)
Dummy for gifts for operating license	Gifts expected or requested to obtain a operating license, conditional on submit a operating license.	1479 (21.6)
Absenteeism	Days of production lost due to absenteeism.	6387 (93.3)

Table A.III: Definitions of Investment climate (IC) variables of Finance and corporate governance group.

Name of the variable	Description of the variable	Observations (Response rate %)
Dummy for commerce chamber	Dummy that takes value 1 if the plant belongs to any association or trade chamber.	6774 (98.9)
Largest shareholder	Percentage of firm's capital owned by the largest shareholder.	6723 (98.2)
Working capital financed by internal funds	Percentage of firm's working capital financed with internal funds.	6477 (94.6)
Working capital financed by domestic commercial banks	Percentage of firm's working capital financed with funds from domestic commercial banks.	6477 (94.6)
Working capital financed by international commercial banks	Percentage of firm's working capital financed with funds from international commercial banks.	6477 (94.6)
Working capital financed by leasing	Percentage of firm's working capital financed with funds from leasing arrangement.	6477 (94.6)
Working capital financed by trade credit	Percentage of firm's working capital financed with credits from suppliers or customers.	6477 (94.6)
Working capital financed by credit cards	Percentage of firm's working capital financed with credit cards.	6477 (94.6)
Working capital financed by equity, sale of stock	Percentage of firm's working capital financed with equity, sale of stock.	6477 (94.6)
Working capital financed by family/friends	Percentage of firm's working capital financed with family/friends funds.	6477 (94.6)
Working capital financed by informal sources	Percentage of firm's working capital financed with funds from informal sources.	6477 (94.6)
New fixed assets financed by internal funds	Percentage of investments in new fixed assets financed with internal funds.	4422 (64.6)
New fixed assets financed by domestic commercial banks	Percentage of investments in new fixed assets financed with funds from domestic commercial banks.	4422 (64.6)
New fixed assets financed by international commercial banks	Percentage of investments in new fixed assets financed with funds from international commercial banks.	4422 (64.6)
New fixed assets financed by leasing	Percentage of investments in new fixed assets financed with funds from leasing arrangement.	4422 (64.6)
New fixed assets financed by trade credit	Percentage of investments in new fixed assets financed with credits from suppliers or customers.	4422 (64.6)
New fixed assets financed by credit cards	Percentage of investments in new fixed assets financed with credit cards.	4422 (64.6)
New fixed assets financed by equity, sole of stock	Percentage of investments in new fixed assets financed with equity, sole of stock.	4422 (64.6)
Financing of new fixed assets financed by family/friends	Percentage of investments in new fixed assets financed with family/friends funds.	4422 (64.6)
New fixed assets financed by informal sources	Percentage of investments in new fixed assets financed with funds from informal sources.	4422 (64.6)
Dummy for current or saving account	Dummy taking value 1 if the plant has a current or saving account.	6810 (99.4)
Dummy for credit line	Dummy that takes value 1 if the firm has access to a credit line or overdraft facility	6834 (99.8)
Dummy for loan	Dummy that takes value 1 if the firm has access to a loan line.	6567 (95.9)
Dummy for loan with collateral	Dummy that takes value 1 if the firm has access to a loan line with collateral (conditional on having a loan line).	2307 (33.7)
Value of the collateral	Value of the collateral as a percentage of the loan value (conditional on having a loan with collateral)	1713 (25)
Duration of the loan	Duration of the loan in months.	2115 (30.9)
Foreign currency	Percentage of total borrowing denominated in foreign currency.	6135 (89.6)
Dummy for external audit	Dummy that takes value 1 if the firm has its annual statements externally audited.	6825 (99.6)

Table A.IV: Definitions of Investment climate (IC) variables of Quality, innovation and labor skills group.

Name of the variable	Description of the variable	Observations (Response rate %)
Dummy for quality certification	Dummy taking value 1 if the firm has any kind of quality certification.	6789 (99.1)
Dummy for foreign technology	Dummy taking value 1 if the plant uses technology licensed from a foreign-owned company.	6549 (95.6)
Dummy for new product	Dummy that takes value 1 if the plant has developed a new product line.	6807 (99.4)
Dummy for product innovation	Dummy taking value 1 if the plant has introduced any product innovation in the last 3 years.	6801 (99.3)
Dummy for joint venture	Dummy that takes value 1 if the firm has established a joint venture with a foreign partner.	6657 (97.2)
Outsourcing	Percentage of total annual sales subcontracted.	6756 (98.6)
Dummy for R&D	Dummy that takes value 1 if the firm performed R&D activities during last year.	6825 (99.6)
Staff – female workers	Percentage of female workers in staff.	5412 (79)
Staff – part-time workers	Percentage of part-time workers in staff.	5415 (79.1)
Staff – primary school	Percentage of workers in staff has not completed primary school.	6147 (89.8)
Staff – skilled workers	Percentage of skilled production workers in staff.	6276 (91.6)
Staff – unskilled workers	Percentage of unskilled production workers in staff.	6099 (89)
Dummy for internal training	Dummy taking value one if the firm provides formal (beyond on the job) internal training to its employees.	6450 (94.2)
Dummy for external training	Dummy taking value one if the firm provides formal (beyond on the job) external training to its employees.	6423 (93.8)
Manager experience	Manager experience in years.	5970 (87.2)
Manager education	Average number of years of education.	5745 (83.9)
Manager tenure	Average tenure in years.	5634 (82.3)
Manager age	Average age.	6000 (87.6)
Workforce with computer	Percentage of workforce using computer at job	6768 (98.8)

Table A.V: Definitions of Investment climate (IC) variables of Other control variables group.

Name of the variable	Description of the variable	Observations (Response rate %)
Age of the firm	Age of the firm in 2005.	6777 (98.9)
Capacity utilization	Percentage of capacity utilized.	6756 (98.6)
Trade union	Percentage of workforce unionized	6396 (93.4)
Strikes	Days of production lost due to strikes.	6327 (92.4)
Dummy for incorporated company	Dummy that takes value 1 if the firm is an incorporated company.	6849 (100)
Dummy for FDI	Dummy that takes value 1 if any part of firm's capital is foreign.	6780 (99)
Number of competitors	Total number of competitors in the domestic market of its major product line.	3186 (46.5)
Dummy for importer	Dummy taking value 1 if the firm imports more than 10% of the total purchases of intermediate materials.	6288 (91.8)
Share of imports	Share of imported inputs over total purchases of intermediate materials.	6288 (91.8)
Dummy for exporter	Dummy taking value 1 if the firm exports more than 10% of the total annual sales.	6312 (92.2)
Share of exports	Share of exports over total annual sales.	6312 (92.2)

Appendix B: Tables and figures

Table B.I: Total number of observations before and after cleaning missing values and outliers.

	Observations before cleaning	Observations after cleaning
Total number of observations in the survey (number of firms multiplied by the number of years)	**6849**	
Missing observations	**2249**	**918**
Of which:		
firms with one PF variable missing	1100	15
firms with two PF variables missing	135	3
firms with three PF variables missing	115	1
firms with four PF variables missing	899	899
Outliers	**241**	**195**
Of which:		
Outliers in materials	126	85
Outliers in labor cost	68	36
Outliers in both materials and labor cost	47	74
Useful observations (outliers and missing excluded) (number of firms multiplied by the number of years)	**4448**	**5750**

The cleaning process is performed in three steps*.

I. Those firms with missing values in all the PF variables (sales, materials, labor cost and capital) are dropped from the sample. For the rest of the missing values we apply the procedure described in II and III.

II. We replace those observations with ratios materials to sales or labor cost to sales greater than one (outliers) following step III.

III. We replace the missing values of the PF variables by their corresponding industry-region-size medians. If we do not have enough observations in some cells, we replace them by the corresponding industry-size medians. If we still do not have enough observations in those cells, in the next step we replace the missing values by the region-size medians. If still necessary, in the last step we compute the medians only by size and/or by industry to replace those missing values.

The last row of the table summarizes the number of useful observations for regression analysis before and after the cleaning process.

*See Escribano and Pena (2009),

Table B.II: Percentage of observations available due to missing values, by industry and size

		Small		Médium		Large		Total	
Industry		**#Obs**	**Perc.**	**#Obs**	**Perc.**	**#Obs**	**Perc.**	**#Obs**	**Perc.**
Food	Original Sample	333		177		87		597	
	Without replacing	177	53.15	79	44.63	51	58.62	307	51.42
	With replacing	248	74.47	137	77.4	69	79.31	454	76.05
Textiles & Leather	Original Sample	426		255		207		888	
	Without replacing	251	58.92	210	82.35	139	67.15	600	67.57
	With replacing	325	76.29	235	92.16	178	85.99	738	83.11
Apparels	Original Sample	360		315		150		825	
	Without replacing	247	68.61	267	84.76	120	80	634	76.85
	With replacing	287	79.72	290	92.06	138	92	715	86.67
Chemicals & Chemical prds	Original Sample	426		333		171		930	
	Without replacing	262	61.5	218	65.47	130	76.02	610	65.59
	With replacing	337	79.11	282	84.68	150	87.72	769	82.69
Plastics & Rubbers	Original Sample	279		189		12		480	
	Without replacing	193	69.18	112	59.26	11	91.67	316	65.83
	With replacing	243	87.1	157	83.07	11	91.67	411	85.62
Non-metallic products	Original Sample	105		63		48		216	
	Without replacing	40	38.1	38	60.32	32	66.67	110	50.93
	With replacing	75	71.43	50	79.37	39	81.25	164	75.93
Structural metal & metal prds	Original Sample	618		252		39		909	
	Without replacing	328	53.07	131	51.98	21	53.85	480	52.81
	With replacing	526	85.11	214	84.92	31	79.49	771	84.82
Machinery & Equipment	Original Sample	1074		687		243		2004	
	Without replacing	749	69.74	482	70.16	160	65.84	1391	69.41
	With replacing	912	84.92	603	87.77	213	87.65	1728	86.23
Total	Original Sample	3621		2271		957		6849	
	Without replacing	2247	62.05	1537	67.68	664	69.38	4448	64.94
	With replacing	2953	81.55	1968	86.66	829	86.62	5750	83.95

Source: Authors' calculations with India ICS data.

Table B.III: Percentage of observations available due to missing values, by industry and year.

Year		2002		2003		2004		Total	
Industry		**#Obs**	**Perc.**	**#Obs**	**Perc.**	**#Obs**	**Perc.**	**#Obs**	**Perc.**
Food	Original Sample	199		199		199		597	
	Without replacing	71	35.68	109	54.77	127	63.82	307	51.42
	With replacing	128	64.32	155	77.89	171	85.93	454	76.05
Textiles & Leather	Original Sample	296		296		296		888	
	Without replacing	164	55.41	212	71.62	224	75.68	600	67.57
	With replacing	218	73.65	256	86.49	264	89.19	738	83.11
Apparels	Original Sample	275		275		275		825	
	Without replacing	193	70.18	214	77.82	227	82.55	634	76.85
	With replacing	222	80.73	240	87.27	253	92	715	86.67
Chemicals & Chemical prds	Original Sample	310		310		310		930	
	Without replacing	175	56.45	214	69.03	221	71.29	610	65.59
	With replacing	233	75.16	265	85.48	271	87.42	769	82.69
Plastics & Rubbers	Original Sample	160		160		160		480	
	Without replacing	88	55	107	66.87	121	75.62	316	65.83
	With replacing	125	78.12	138	86.25	148	92.5	411	85.62
Non-metallic products	Original Sample	72		72		72		216	
	Without replacing	26	36.11	41	56.94	43	59.72	110	50.93
	With replacing	44	61.11	59	81.94	61	84.72	164	75.93
Structural metal & metal prds	Original Sample	303		303		303		909	
	Without replacing	128	42.24	162	53.47	190	62.71	480	52.81
	With replacing	231	76.24	264	87.13	276	91.09	771	84.82
Machinery & Equipment	Original Sample	668		668		668		2004	
	Without replacing	411	61.53	470	70.36	510	76.35	1391	69.41
	With replacing	531	79.49	588	88.02	609	91.17	1728	86.23
Total	Original Sample	2283		2283		2283		6849	
	Without replacing	1256	55.02	1529	66.97	1663	72.84	4448	64.94
	With replacing	1732	75.87	1965	86.07	2053	89.93	5750	83.95

Source: Authors' calculations with India ICS data.

Table C.I: Robust IC elasticities and semi-elasticities with respect to productivity – OLS Estimation.

Blocks of ICA variables	Explanatory ICA variables	Two steps		Single step estimation			
		Solow residual		Cobb-Douglas		Translog	
		Restricted	Unrestric.	Restricted	Unrestric.	Restricted	Unrestric.
Infrastructures	Longest days to clear customs to export (a)	-0.016	-0.032	-0.01	-0.013	-0.015	-0.013
	Dummy for own generator	0.062**	0.066**	0.057*	0.081***	0.056*	0.095***
	Water supply from public sources (b)	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	Shipment losses in the domestic market (b)	-0.004*	-0.003	-0.005*	-0.005**	-0.005**	-0.005**
	Dummy for own transport	0.035	0.03	0.022	0.012	0.036	0.052
	Dummy for web page	0.036	0.035	0.039	0.054**	0.038	0.044*
Red Tape, informality and others	Dummy for security (b)	0.051**	0.052**	0.045*	0.059**	0.044*	0.058***
	Sales reported to taxes (b)	0.001	0	0.001	0.001	0.001	0.001
	Workforce reported to taxes (b)	-0.002**	-0.001	-0.002**	-0.002**	-0.002**	-0.002***
	Dummy for payments to speed up bureaucracy	-0.038*	-0.053**	-0.044**	-0.041*	-0.048**	-0.048**
	Dummy for interventionist labor regulation	-0.038	-0.029	-0.036	-0.050*	-0.024	-0.019
	Absenteeism (b)	-0.032**	-0.045***	-0.030**	-0.036***	-0.033**	-0.036***
Finance and Corporate Governance.	Dummy for trade association	0.065**	0.077***	0.074***	0.068**	0.084***	0.099***
	Working capital financed by domestic private banks (b)	0	0	0	0	0	0
	Dummy for loan (b)	0.063**	0.069**	0.068**	0.078***	0.073***	0.072***
	Dummy for external audit.	0.119***	0.132***	0.116***	0.123***	0.109***	0.116***
Quality, Innovation and labor skills	Dummy for R&D (a)	0.15	0.341**	0.148	0.196	0.133	0.058
	Dummy for product innovation	-0.003	0.013	-0.005	-0.011	-0.003	-0.001
	Dummy for foreign license (b)	0.188***	0.169***	0.200***	0.154**	0.176***	0.155***
	Dummy for internal training (b)	0.059	0.068*	0.059	0.056	0.05	0.087***
	Workforce with computer	0.002**	0.002**	0.002**	0.002***	0.002**	0.002***
Other control variables	Dummy for incorporated company	0.037	0.045*	0.027	0.043	0.017	0.046*
	Age of the firm	0.059***	0.056***	0.055***	0.040**	0.056***	0.043***
	Share of exports (b)	0.001*	0.001	0.001	0.001	0.001	0.001
	Trade union (b)	0.001	0	0.001	0.001	0.001	0.001
	Strikes (b)	-0.067**	-0.094**	-0.068**	-0.073**	-0.051*	-0.012
	Observations	5230	5230	5230	5230	5230	5230
	R-squared	0.084	0.086	0.881	0.89	0.884	0.903

NOTES:
Two steps estimation: in the first step estimation of equation (b2.1) by non-parametric techniques to compute productivity (Solow residual), in the second step estimate (3.2) and (3.3) by OLS using as dependent variable the Solow residual from the first step, either restricted or unrestricted.
Single step estimation: estimate (3.1), (3.2) and (3.3) in a single step by OLS, where (3.1) can be a Cobb-Douglas Production function or a Translogarithmic.
Restricted: equal input output for all the establishments in the country.
Unrestricted: equal input-output elasticities for all the establishments in the same sector.
*significant at 10%; ** significant at 5%; *** significant at 1% given by robust standard errors corrected for correlation between cluster (industry and region).
Each regression includes a set of industry, size and region dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.II: Further robustness; IC elasticities and semi-elasticities with respect to productivity – Random effects estimation.

Blocks of ICA variables		Two steps		Single step estimation			
		Solow residual		Cobb-Douglas		Translog	
	Explanatory ICA variables	Restricted	Unrestric.	Restricted	Unrestric.	Restricted	Unrestric.
Infrastructures	Longest days to clear customs to export (a)	-0.01	-0.025	-0.013	-0.021	-0.031	-0.008
	Dummy for own generator	0.063	0.067	0.143***	0.165***	0.131***	0.162***
	Water supply from public sources (b)	0.002***	0.001***	0.001**	0.001**	0.002***	0.001***
	Shipment losses in the domestic market (b)	-0.004	-0.003	-0.006	-0.006	-0.007	-0.007
	Dummy for own transport	0.052	0.044	0.106*	0.09	0.133**	0.146***
	Dummy for web page	0.029	0.029	0.075	0.081*	0.063	0.042
Red Tape, informality and others	Dummy for security (b)	0.043	0.043	0.107**	0.122***	0.095**	0.082**
	Sales reported to taxes (b)	0.001	0	0.001	0.002	0.002	0.002
	Workforce reported to taxes (b)	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002**
	Dummy for payments to speed up bureaucracy	-0.05	-0.065*	-0.05	-0.037	-0.06	-0.053
	Dummy for interventionist labor regulation	-0.043	-0.035	-0.032	-0.046	-0.013	0.023
	Absenteeism (b)	-0.021	-0.032	-0.032	-0.041*	-0.043*	-0.044**
Finance and Corporate Governance.	Dummy for trade association	0.066	0.076*	0.08	0.079*	0.105**	0.139***
	Working capital financed by domestic private banks (b)	0.001	0	0.002**	0.002**	0.001**	0.001
	Dummy for loan (b)	0.059	0.066*	0.074*	0.085**	0.086**	0.098**
	Dummy for external audit.	0.124**	0.137**	0.155**	0.163***	0.135**	0.150***
Quality, Innovation and labor skills	Dummy for R&D (a)	0.137	0.317*	0.218	0.274	0.334*	0.405**
	Dummy for product innovation	0.004	0.02	0.021	0.016	0.011	0.005
	Dummy for foreign license (b)	0.216***	0.194**	0.280***	0.205**	0.224**	0.200**
	Dummy for internal training (b)	0.046	0.057	0.063	0.046	0.021	0.078
	Workforce with computer	0.002**	0.002**	0.003***	0.004***	0.003***	0.004***
Other control variables	Dummy for incorporated company	0.033	0.041	0.129***	0.133***	0.093*	0.091**
	Age of the firm	0.057**	0.054**	0.081***	0.061**	0.081***	0.055**
	Share of exports (b)	0.001**	0.001*	0.003***	0.002***	0.003***	0.002***
	Trade union (b)	0.001	0.001	0.002*	0.001	0.002*	0.002**
	Strikes (b)	-0.079*	-0.104**	-0.042	-0.07	-0.004	0.003
	Observations	5230	5230	5230	5230	5230	5230
	R-squared	0.084	0.086	0.881	0.89	0.884	0.903

NOTES:

Two steps estimation: in the first step estimation of equation (b2.1) by non-parametric techniques to compute productivity (Solow residual), in the second step estimate (3.2) and (3.3) by OLS using as dependent variable the Solow residual from the first step, either restricted or unrestricted.

Single step estimation: estimate (3.1), (3.2) and (3.3) in a single step by OLS, where (3.1) can be a Cobb-Douglas Production function or a Translogarithmic.

Restricted: equal input output for all the establishments in the country.

Unrestricted: equal input-output elasticities for all the establishments in the same sector.

*significant at 10%; ** significant at 5%; *** significant at 1% given by robust standard errors corrected for correlation between cluster (industry and region).

Each regression includes a set of industry, size and region dummies and a constant term.

(a) Variables instrumented with the industry-region-size average.

(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).

Source: Authors' calculations with India ICS data.

Table C.III: IC percentage contributions to aggregate log-productivity.

			Aggregate log-TFP	Average log-TFP	Allocative efficiency
Demean log-productivity	**Infrastructures**	Longest days to clear customs to export (a)	-1.27	-1.38	0.11
		Dummy for own generator	1.65	1.00	0.65
		Water supply from public sources (b)	1.65	1.06	0.59
		Shipment losses in the domestic market (b)	-0.09	-0.13	0.04
		Dummy for own transport	0.32	0.13	0.18
		Dummy for web page	0.52	0.40	0.12
	Red Tape, informality and others	Dummy for security (b)	1.46	1.10	0.36
		Sales reported to taxes (b)	2.01	2.03	-0.01
		Workforce reported to taxes (b)	-4.27	-4.18	-0.08
		Dummy for payments to speed up bureaucracy	-0.56	-0.66	0.10
		Dummy for interventionist labor regulation	-1.08	-1.00	-0.08
		Absenteeism (b)	-0.30	-0.31	0.01
	Finance and Corporate Governance.	Dummy for trade association	2.06	1.74	0.32
		Working capital financed by domestic private banks (b)	0.59	0.39	0.19
		Dummy for loan (b)	1.20	0.74	0.46
		Dummy for external audit.	3.88	3.33	0.55
	Quality, Innovation and labor skills	Dummy for R&D (a)	1.95	1.42	0.53
		Dummy for product innovation	-0.13	-0.11	-0.02
		Dummy for foreign license (b)	1.15	0.31	0.84
		Dummy for internal training (b)	0.56	0.26	0.30
		Workforce with computer	1.09	0.84	0.26
	Other control variables	Dummy for incorporated company	0.94	0.40	0.53
		Age of the firm	5.78	5.29	0.49
		Share of exports (b)	0.87	0.47	0.40
		Trade union (b)	0.60	0.11	0.48
		Strikes (b)	-0.32	-0.13	-0.20
Total contribution of IC (demean log-productivity)			20.26	13.12	7.13
Other stuff		Industry/region/size controls	-1.09	-4.89	3.80
		Constant term	23.05	23.05	0.00
		Residual	57.78	0.00	57.78
Total contribution of other stuff			79.74	18.16	61.59
Total			100.00	31.28	68.72

NOTES:
Results from equation (5.3)..
Demeaned log-productivity is the part of productivity associated with the investment climate
The productivity measure used is the restricted Solow residual.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.IV: Extended production function and comparison of ICA method with EM algorithms.

Dependent variable: Log of total sales		ICA Method[1]			EM Algorithms[2] [1]		[2]		[3]	
Category	Variable	Coeff.	std. err.	Boot. s.e	Coeff.	std. err.	Coeff.	std. err.	Coeff.	std. err.
PF variables	Log-employment	0.1027	[0.0341]***	(0.0306)***	0.0976	[0.0331]***	0.0516	[0.0250]**	0.0527	[0.0250]**
	Log-materials	0.7989	[0.0185]***	(0.0462)***	0.8362	[0.0186]***	0.8607	[0.0176]***	0.8628	[0.0177]***
	Log-capital	0.0676	[0.0239]***	(0.0153)***	0.0629	[0.0225]***	0.0537	[0.0146]***	0.0502	[0.0147]***
Infrastructures	Longest # of days to clear customs for exports (a)	-0.0125	[0.0263]	(0.0376)	-0.0039	[0.0275]	-0.0158	[0.0209]	-0.0156	[0.0208]
	Dummy for own generator	0.0538	[0.0422]	(0.0424)	0.0378	[0.0396]	0.015	[0.0247]	0.0131	[0.0249]
	Water supply from public sources (b)	0.0014	[0.0005]***	(0.0008)*	0.0013	[0.0004]***	0.0009	[0.0003]***	0.0008	[0.0003]**
	Shipment losses in the domestic market (b)	-0.0047	[0.0039]	(0.0128)	-0.0023	[0.0035]	-0.0017	[0.0030]	-0.0016	[0.0030]
	Dummy for own transport	0.0238	[0.0475]	(0.0861)	-0.0084	[0.0464]	-0.003	[0.0340]	-0.0023	[0.0341]
	Dummy for web page	0.0402	[0.0394]	(0.0264)	0.0047	[0.0378]	0.0013	[0.0310]	0.0008	[0.0313]
Red Tape, informality and others	Dummy for security	0.0467	[0.0423]	(0.1407)	0.0426	[0.0403]	0.0497	[0.0285]*	0.0505	[0.0285]*
	Sales reported for taxes (b)	0.0006	[0.0014]	(0.0052)	0.0009	[0.0013]	0.0008	[0.0010]	0.0009	[0.0010]
	Workforce reported for taxes (b)	-0.0015	[0.0012]	(0.0042)	-0.0015	[0.0010]	-0.0009	[0.0008]	-0.0009	[0.0008]
	Dummy for payments to speed up bureaucracy	-0.0464	[0.0336]	(0.0526)	-0.0443	[0.0292]	0.0041	[0.0255]	0.0083	[0.0259]
	Dummy for interventionist labor regulation	-0.036	[0.0361]	(0.0211)*	-0.0317	[0.0340]	-0.0259	[0.0330]	-0.028	[0.0331]
	Absenteeism (b)	-0.0299	[0.0222]	(0.0571)	-0.0204	[0.0195]	-0.0069	[0.0156]	-0.0071	[0.0160]
Finance and corporate governance	Dummy for trade association	0.0785	[0.0455]*	(0.0456)*	0.0756	[0.0408]*	0.024	[0.0297]	0.0194	[0.0300]
	Working capital financed by domestic private banks (b)	0.0002	[0.0007]	(0.0005)	-0.0002	[0.0007]	0.0003	[0.0006]	0.0003	[0.0006]
	Dummy for external audit	0.0691	[0.0395]*	(0.0452)	0.0662	[0.0362]*	0.0633	[0.0283]**	0.0655	[0.0282]**
	Dummy for loan (b)	0.1102	[0.0473]**	(0.0637)*	0.0892	[0.0464]*	0.0121	[0.0331]	0.006	[0.0327]
Quality, innovation and labor skills	Dummy for R&D (a)	0.1787	[0.2382]	(0.2347)	0.2041	[0.2534]	0.0702	[0.1322]	0.0638	[0.1320]
	Dummy for product innovation	-0.0073	[0.0360]	(0.0710)	-0.0153	[0.0332]	-0.025	[0.0244]	-0.0265	[0.0246]
	Dummy for foreign license (b)	0.204	[0.1053]*	(0.1302)	0.1425	[0.1033]	0.086	[0.0847]	0.0801	[0.0852]
	Dummy for internal training (b)	0.0579	[0.0533]	(0.0516)	0.0578	[0.0511]	0.0702	[0.0443]	0.0703	[0.0442]
	Unskilled workforce (a)	0.0013	[0.0036]	(0.0016)	0.0013	[0.0036]	-0.0034	[0.0030]	-0.0039	[0.0031]
	Workforce with computer	0.0017	[0.0011]	(0.0015)	0.0016	[0.0010]	0.0012	[0.0009]	0.0011	[0.0008]
Other control variables	Dummy for incorporated company	0.0265	[0.0396]	(0.0901)	0.0162	[0.0368]	0.0272	[0.0301]	0.0261	[0.0300]
	Age of the firm	0.0534	[0.0267]**	(0.0214)**	0.0438	[0.0251]*	0.0456	[0.0174]**	0.0487	[0.0174]***
	Share of exports (b)	0.001	[0.0009]	(0.0005)**	0.0006	[0.0009]	0.00004	[0.0006]	-0.0001	[0.0006]
	Trade union (b)	0.0008	[0.0012]	(0.0010)	0.0008	[0.0012]	0.0009	[0.0009]	0.0007	[0.0009]
	Strikes (b)	-0.0683	[0.0449]	(0.0821)	-0.0475	[0.0380]	-0.0112	[0.0307]	-0.0107	[0.0314]
	Constant	0.7377	[0.3449]**		0.4456	[0.3504]	1.0108	[0.2499]***	1.0335	[0.2492]***
	Industry/region/size/time dummies	Yes			Yes		Yes		Yes	
	Observations	5211			5216		5175		5176	
	R-squared	0.88			0.9		0.94		0.94	

Estimating results of equation (1) under different imputation mechanisms for missing data. Those observations with missing values in all sales, labor (labor cost), materials and capital are excluded in all the regressions.
[1] ICA method is in section 3 of main text. Significance is given by clustered and White-robust standard errors in brackets; *** 1%, **5%, * 10%. In parentheses are bootstrap standard errors after 1000 replications (see section 5.2.2 on the motivation of using bootstrap standard errors). Correlation by clusters is also considered.
[2] EM algorithms are explained in section 5.1. EM alg [1] includes as covariates of the imputation mechanism industry/region/size/time (I/R/S/T) dummies (see section 5.1.1); EM alg [2] includes I/R/S/T dummies and production function variables (see section 5.1.2); EM alg [3] also includes a set of IC variables (see section 5.1.3). Significance is given by clustered White-robust standard errors. (a) IC variables instrumented with industry/region average variables. (b) missing values in IC variables replaced by means of *ICA method*.
Source: Escribano and Pena (2010).

Table D.I: IC elasticities and semi-elasticities with respect to employment – IV Estimation.

Dependent variable: employment		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.447**	14.8	0.405**	13.1
Real wages[1]		-0.431***	-97.2	-0.316**	-71.3
Infrastructures	Dummy for own generator (b)	0.313***	5.2	0.304***	5.1
	Shipment losses in the domestic market (b)	-0.020***	-0.8	-0.018***	-0.7
	Dummy for own transport (b)	0.178***	0.7	0.159***	0.6
	Dummy for web page (b)	0.312***	3.1	0.311***	3.1
Red Tape, informality and others	Sales reported to taxes (b)	0.003***	9.4	0.003***	8.1
	Dummy for security (b)	0.205***	4.3	0.209***	4.4
Finance and corporate governance	Largest shareholder (b)	-0.005***	-12.5	-0.005***	-11.7
	Dummy for chamber of commerce(b)	0.079*	2.1	0.090**	2.4
	Dummy for overdraft (b)	0.154***	2.7	0.162***	2.9
	Borrowing denominated in foreign currency (b)	0.006**	0.3	0.006**	0.3
Quality, innovation and labor skills	Dummy for quality certification (b)	0.326***	2.1	0.322***	2.1
	Dummy for R&D (b)	0.313***	2.5	0.289***	2.3
	Staff – female workers (b)	0.003***	0.6	0.003***	0.6
	Staff - skilled workers (a)	0.018***	33.4	0.017***	32.0
	Dummy for internal training (b)	0.229***	1.0	0.219***	1.0
	Workforce with computer (b)	0.003**	1.2	0.003***	1.3
Other control variables	Dummy for incorporate company	0.508***	4.8	0.498***	4.7
	Age of the firm	0.148***	13.9	0.149***	14.0
	Share of exports (b)	0.006***	1.9	0.006***	1.9
	Trade union (a)	0.017***	4.0	0.017***	4.0
Instruments evaluation	First stage R-squared: productivity[2]	0.084		0.082	
	Partial R-squared: productivity[3]	0.109		0.013	
	Partial R-squared F test (p-value): productivity[4]	0.000		0.000	
	First stage R-squared: wages[2]	0.164		0.164	
	Partial R-squared: wages[3]	0.005		0.007	
	Partial R-squared F test (p-value): wages[4]	0.000		0.000	
	Hansen test (p-value)[5]	0.345		0.308	
	Observations	5535		5535	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity and real wages are endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with India ICS data.

Table D.II: IC elasticities and semi-elasticities with respect to real wages – IV Estimation

Dependent variable: real wages		**Restricted Solow residual**		**Unrestricted Solow residual**	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.398*	5.9	0.511*	7.5
Infrastructures	Dummy for own generator (b)	0.093*	0.8	0.079	0.6
	Dummy for own transport (b)	0.223***	0.4	0.216***	0.4
	Shipment losses in the domestic market (b)	-0.017***	-0.3	-0.017***	-0.3
Red Tape, informality and others	Sales declared to taxes (b)	0.002***	3.2	0.002***	3.0
	Dummy for criminal activity (b)	-0.133*	-0.1	-0.130*	-0.1
	Dummy for illegal payments for protection (b)	-0.129	-0.2	-0.144*	-0.2
	Manager's time in bureaucratic issues (b)	0.004**	0.8	0.004**	0.8
Finance and corporate governance	Largest shareholder (b)	-0.002***	-2.7	-0.002***	-2.6
	New fixed assets financed by domestic commercial banks (a)	0.004**	1.9	0.004**	2.0
	Dummy for loan (b)	0.094**	0.5	0.085*	0.5
Quality, innovation and labor skills	Dummy for R&D (b)	0.226***	1.0	0.221***	0.9
	Staff - skilled workers (b)	0.004***	3.1	0.004***	3.1
	Dummy for external training (b)	0.237***	0.3	0.235***	0.3
	Manager tenure (b)	0.095***	3.6	0.092**	3.5
Other control variables	Dummy for incorporate company	0.103*	0.5	0.09	0.5
	Number of competitors (a)	-0.094***	-3.7	-0.096***	-3.8
	Share of imports (b)	0.005***	0.2	0.005**	0.2
Instruments evaluation	First stage R-squared[2]	0.077		0.074	
	Partial R-squared[3]	0.011		0.009	
	Partial R-squared F test (p-value)[4]	0		0	
	Hansen test (p-value)[5]	0.529		0.696	
	Observations	5535		5535	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with India ICS data.

Table D.III: IC linear probability coefficients with respect to the probability of exporting – IV Estimation

Dependent variable: probability of exporting		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.217***	202.9	0.172**	156.4
Infrastructures	Dummy for own generator (b)	0.044***	19.2	0.049***	21.1
	Dummy for own transport (b)	-0.083***	-8.1	-0.078***	-7.7
	Dummy for e-mail (b)	0.037***	18.8	0.041***	20.8
	Dummy for web page (b)	0.082***	20.0	0.084***	20.5
Red Tape, informality and others	Dummy for security cost(b)	0.032***	18.0	0.034***	19.0
	Dummy for interventionist labor regulation (a)	-0.258***	-218.8	-0.282***	-239.4
Finance and corporate governance	Working capital financed by domestic private banks (b)	0.001***	31.8	0.001***	28.1
	Dummy for current account (b)	0.054***	52.5	0.051***	50.0
	Borrowing denominated in foreign currency (a)	0.015***	22.2	0.015***	21.2
Quality, innovation and labor skills	Dummy for R&D (b)	0.030**	6.0	0.029**	5.8
	Dummy for outsourcing (b)	0.041***	6.8	0.044***	7.3
	Dummy for external training (b)	0.066***	2.9	0.064***	2.8
Other control variables	Age of the firm	-0.024***	-63.4	-0.019**	-51.9
	Number of competitors (a)	0.026***	65.4	0.023**	58.2
Instruments evaluation	First stage R-squared[2]	0.073		0.07	
	Partial R-squared[3]	0.01		0.01	
	Partial R-squared F test (p-value)[4]	0		0	
	Hansen test (p-value)[5]	0.194		0.107	
	Observations	5534		5534	

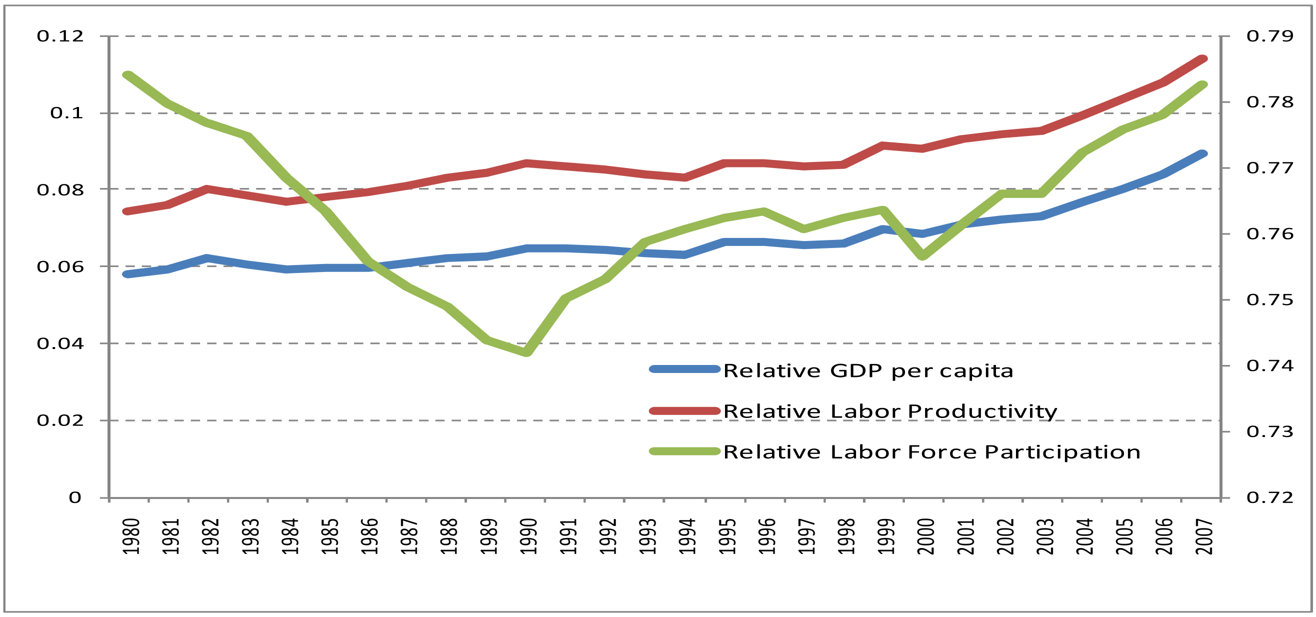
NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with India ICS data.

Table D.IV: IC linear probability coefficients with respect to the probability of receiving FDI – IV Estimation.

Dependent variable: probability of receiving fdi		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.041*	270.1	0.041*	263.5
Infrastructures	Number of power outages (b)	-0.003**	-69.7	-0.003**	-68.1
	Dummy for own generator (b)	0.008*	25.4	0.008*	23.8
	Dummy for own transport (b)	-0.019***	-13.8	-0.019***	-13.6
Red Tape, informality and others	Dummy for conflicts with court involved (b)	0.049***	35.6	0.050***	36.0
	Dummy for criminal activity (b)	-0.024***	-11.3	-0.025***	-11.5
	Dummy for illegal payments for protection (b)	-0.021***	-12.3	-0.022***	-12.7
Finance and corporate governance	Dummy for commercial chamber (b)	0.008**	42.4	0.007**	38.8
	Working capital financed by domestic private banks (b)	0.000***	91.3	0.000***	89.4
	Dummy for overdraft (b)	0.016***	53.0	0.016***	53.7
Quality, innovation and labor skills	Dummy for foreign license (b)	0.065***	19.6	0.066***	19.9
	Unskilled workforce (b)	-0.000***	-65.1	-0.000***	-60.1
	Dummy for external training (b)	0.074***	24.6	0.074***	24.5
	Experience of the manager (b)	0.008***	109.1	0.008***	106.2
	Workforce with computer (b)	0.001***	53.9	0.001***	54.1
Other control variables	Age of the firm	-0.006*	-112.1	-0.006*	-107.7
	Strikes (b)	-0.020***	-6.5	-0.019***	-6.1
Instruments evaluation	First stage R-squared[2]	0.079		0.08	
	Partial R-squared[3]	0.01		0.01	
	Partial R-squared F test (p-value)[4]	0		0	
	Hansen test (p-value)[5]	0.331		0.312	
	Observations	5623		5623	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with India ICS data.

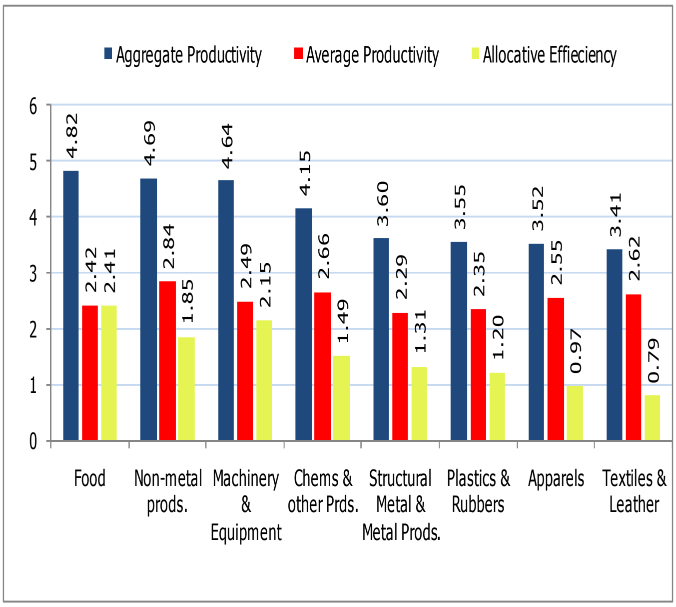
Figure 2: Decomposition of GDP gap between India and USA, 1980/2007



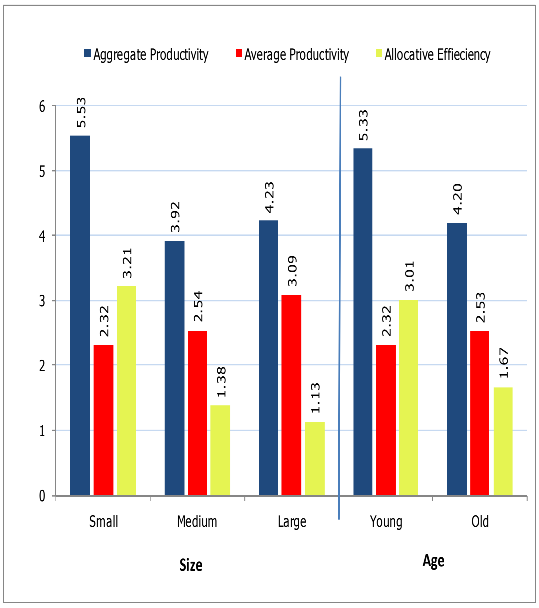
Source: Authors `Calculations with Penn World Table Version 6.3, Center for International Comparisons at the University of Pennsylvania.

Figure 3: Olley and Pakes decomposition in levels.

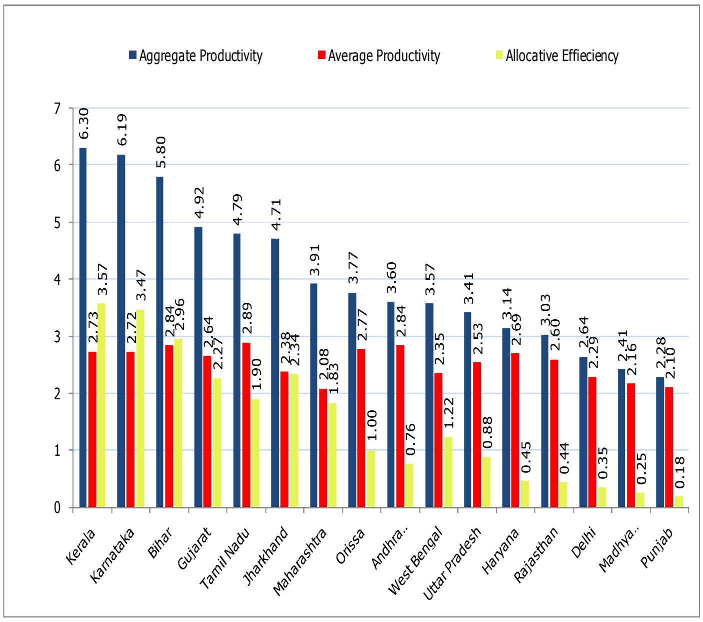
a) by Industry



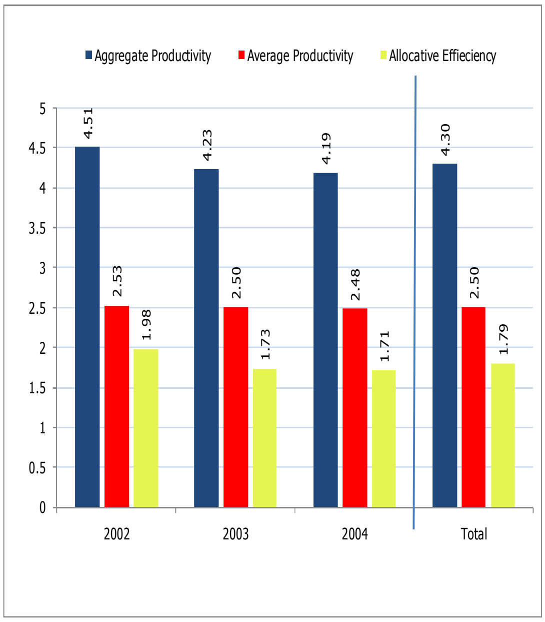
b) by Size and Age



c) by State



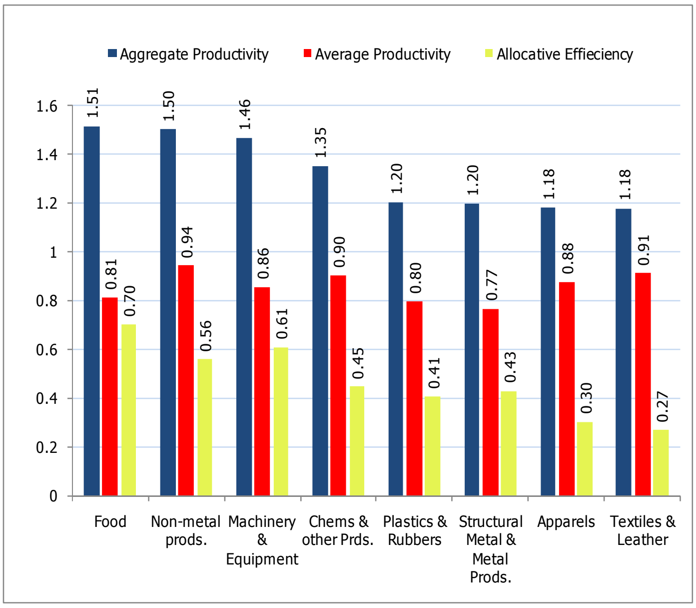
d) by Year



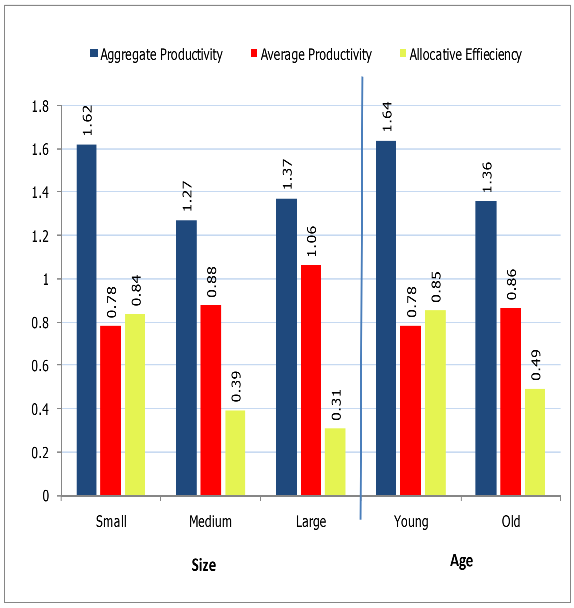
Note: Olley and Pakes decomposition in levels according to equation (5.1a). The productivity measure used is the restricted Solow residual in levels.
Source: Authors' calculations with India ICS data.

Figure 4: *Mixed* Olley and Pakes decomposition.

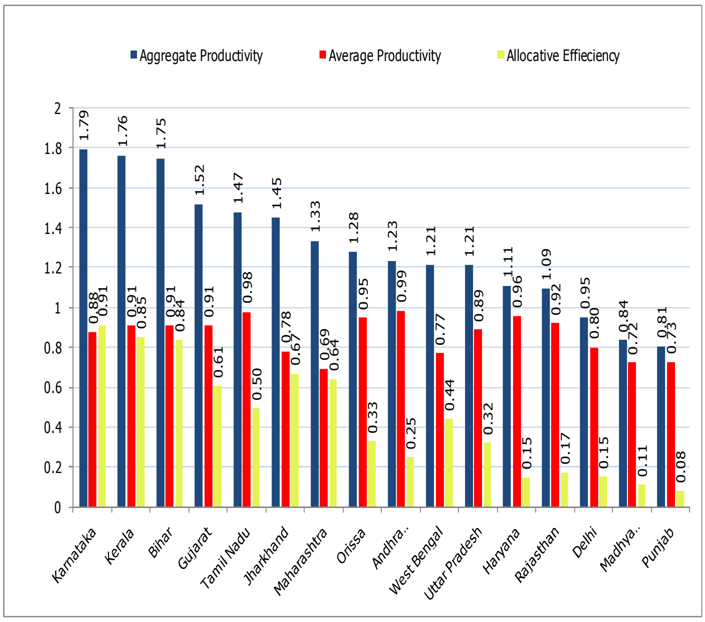
a) by Industry



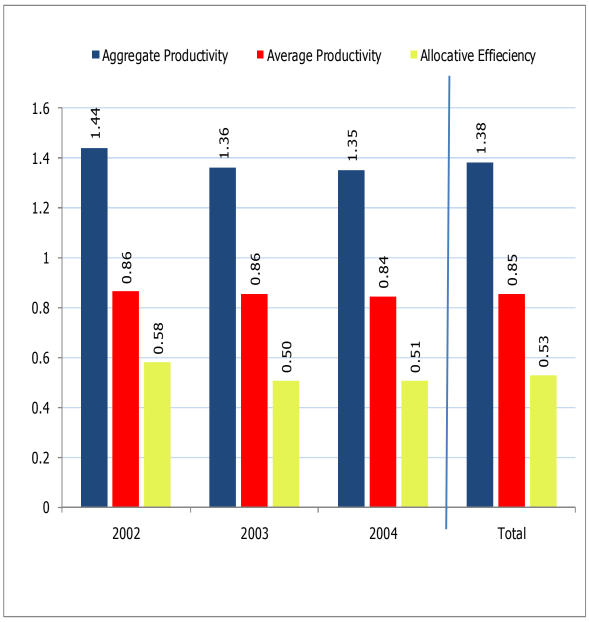
b) by Size and Age



c) by State

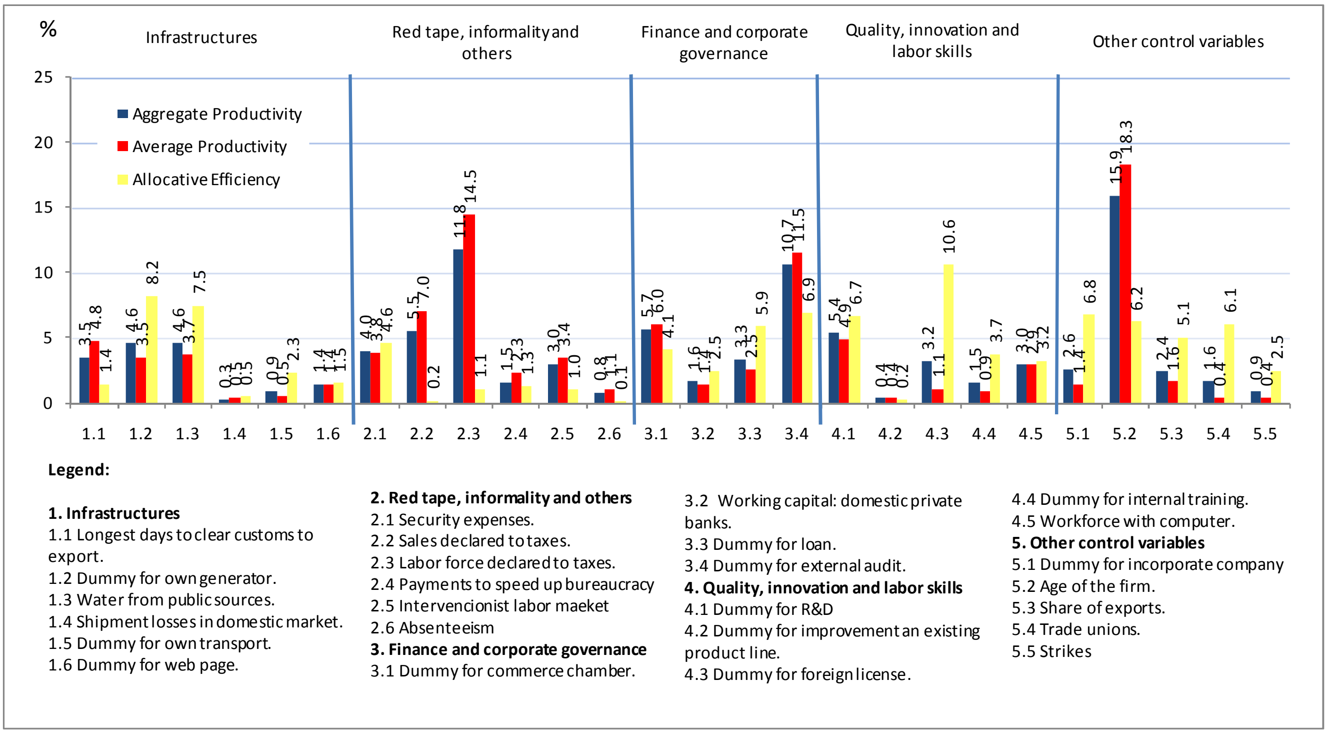


d) by Year



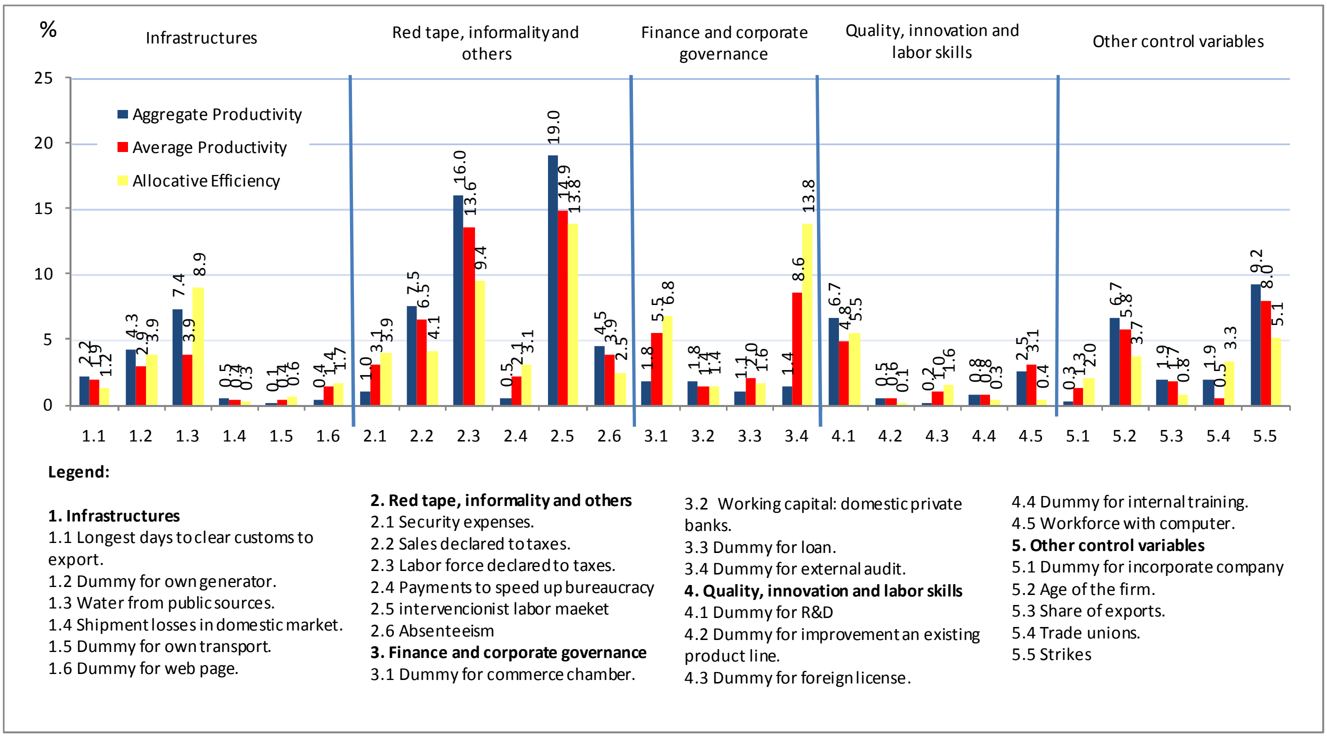
Note: Mixed Olley and Pakes decomposition according to equation (5.1b). The productivity measure used is the restricted Solow residual in logs.
Source: Authors' calculations with India ICS data.

Figure 4.1: IC Percentage contributions to aggregate log-productivity



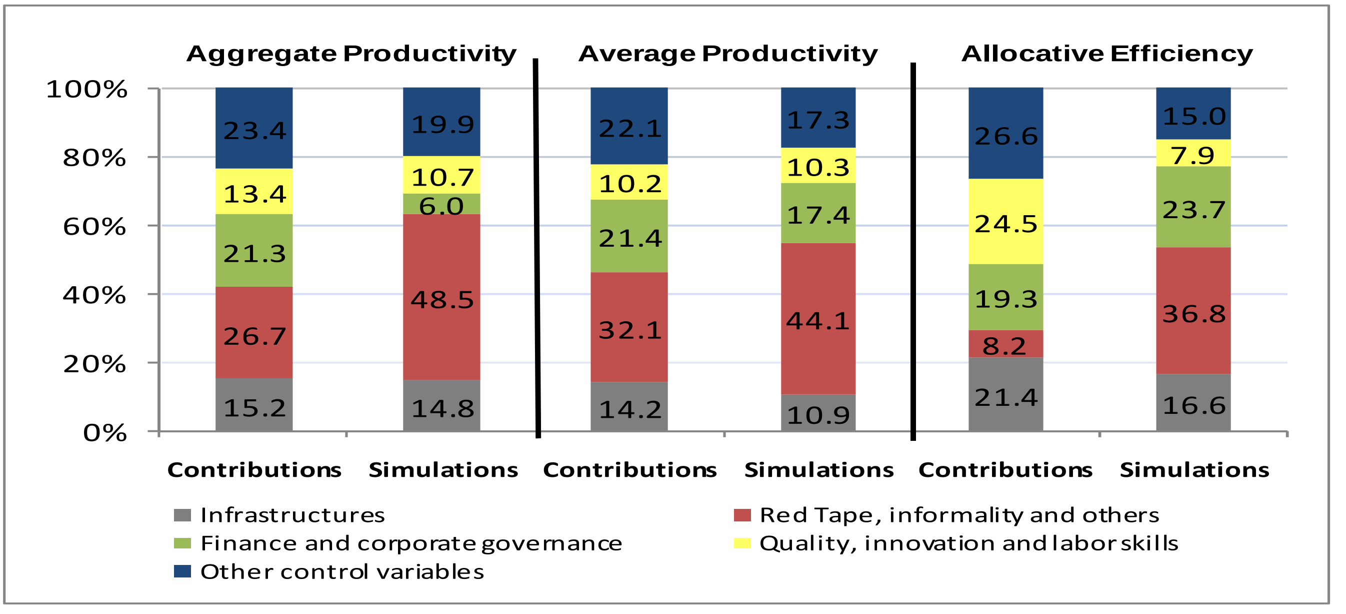
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 4.2: Percentage change in aggregate productivity (TFP) from a 20% improvement of IC variables.



Note: Simulations computed according to section 4.3. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

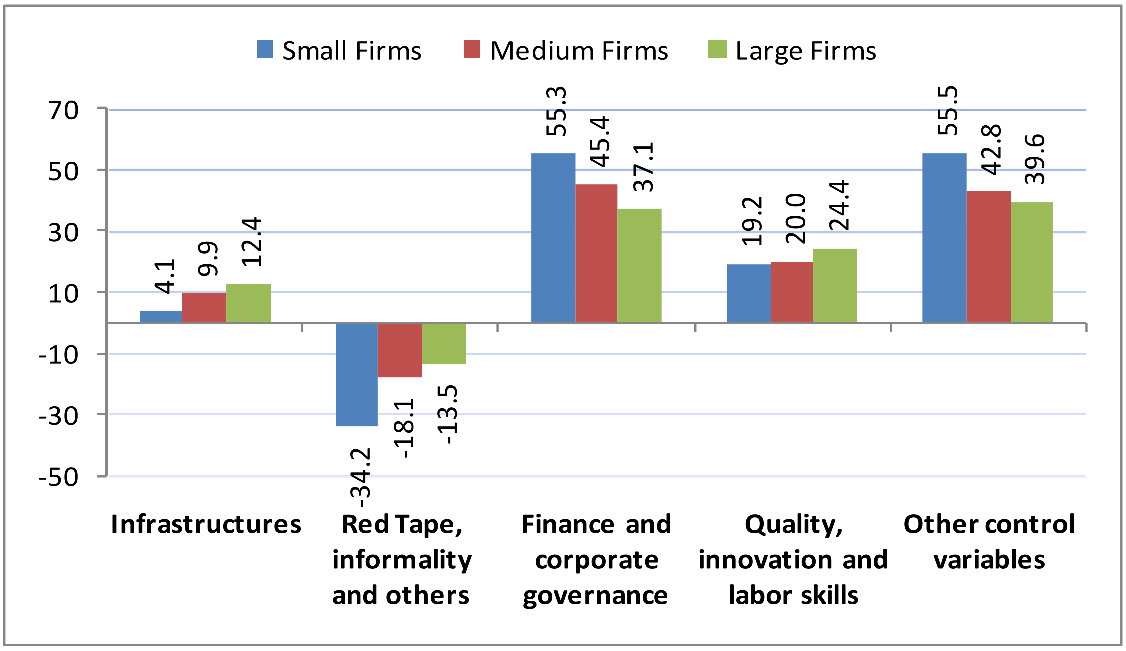
Figure 4.3: Weight of each block of IC variables on aggregate productivity, average productivity and allocative efficiency, by contributions and by simulations.



Note: The weight of each block or group of IC variables from contributions comes from Figure 4.1. We take the percentage contributions of Figure 4.1 in absolute value and we compute the relative weight of each block.
For the case of simulations we do the same with the percentage increases of Figure 4.2.
The productivity measure used is the *demeaned* restricted Solow residual in logs.
Source: Authors' calculations with India ICS data.

Figure 4.4: Contributions of each block of IC variables on average productivity by size.

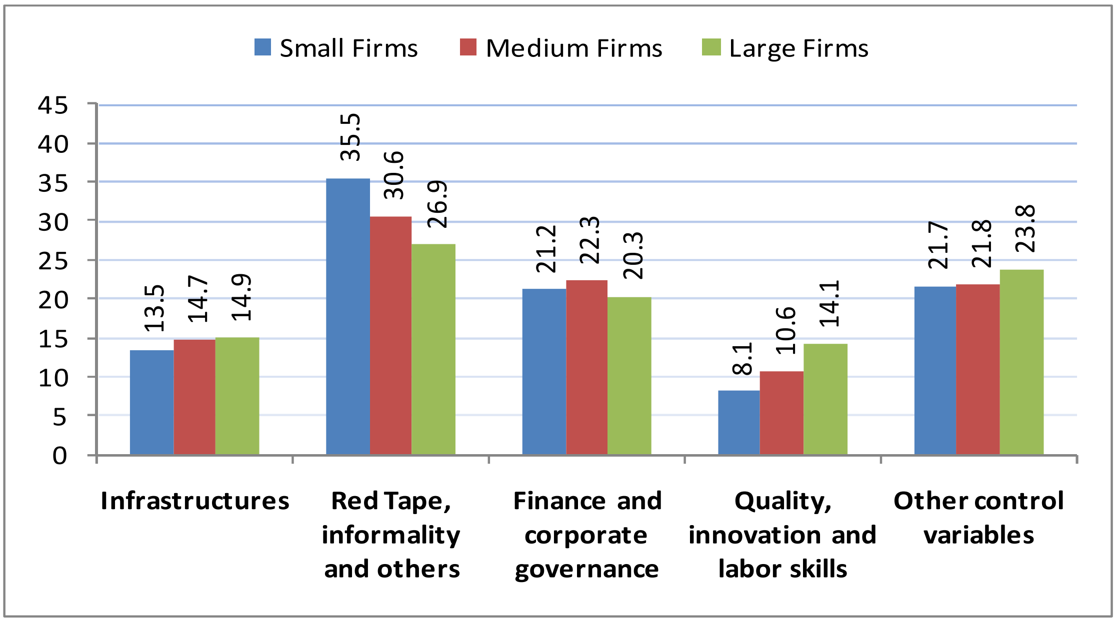
a) percentage contributions



Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.

Source: Authors' calculations with India ICS data.

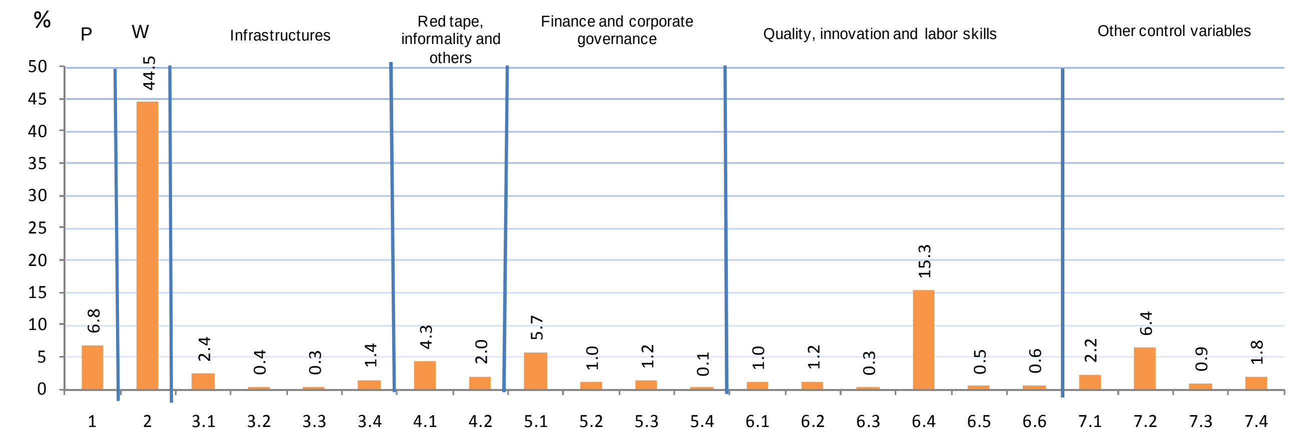
b) absolute percentage contributions



Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.

Source: Authors' calculations with India ICS data.

Figure 5.1: IC percentage contributions to average log-employment.



Legend:

1. Log of Productivity
2. Log of Real Wage

3. Infrastructures
3.1 Dummy for own generator.
3.2 Shipment losses in domestic market.
3.3 Dummy for own transport.
3.4 Dummy for web page.

4. Red tape, informality and others
4.1 Sales declared to taxes.
4.2 Dummy for sucurity.

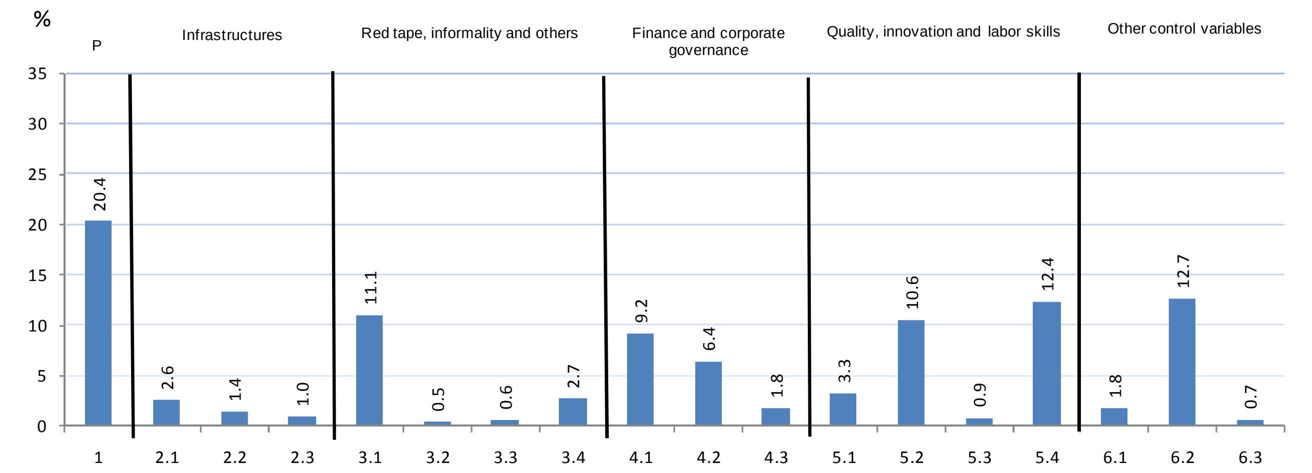
5. Finance and corporate governance
5.1 Largest shareholder
5.2 Dummy for commerce chamber.
5.3 Dummy for overdraft.
5.4 Borrowing denominated in foreign currency.

6. Quality, innovation and labor skills
6.1 Dummy for quality certification.
6.2 Dummy for R&D.
6.3 Staff - female workers.
6.4 Staff - skilled workers.
6.5 Dummy for internal training.
6.6 Workforce with computer.

7. Other control variables
7.1 Dummy for incorporate company
7.2 Age of the firm.
7.3 Share of exports.
7.4 Trade unions.

Note: Contributions computed according to section 5.5. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.2: IC percentage contributions to average log-wage.



Legend:

1. Log of Productivity

2. Infrastructures
2.1 Dummy for own generator.
2.2 Dummy for own transport.
2.3 Shipment losses in the domestic market.

3. Red tape, informality and others
3.1 Sales declared to taxes.
3.1 Dummy for criminal activity.
3.3 Dummy for illegal payments for protection.
3.4 Manager`s time spent in bureaucratic issues.

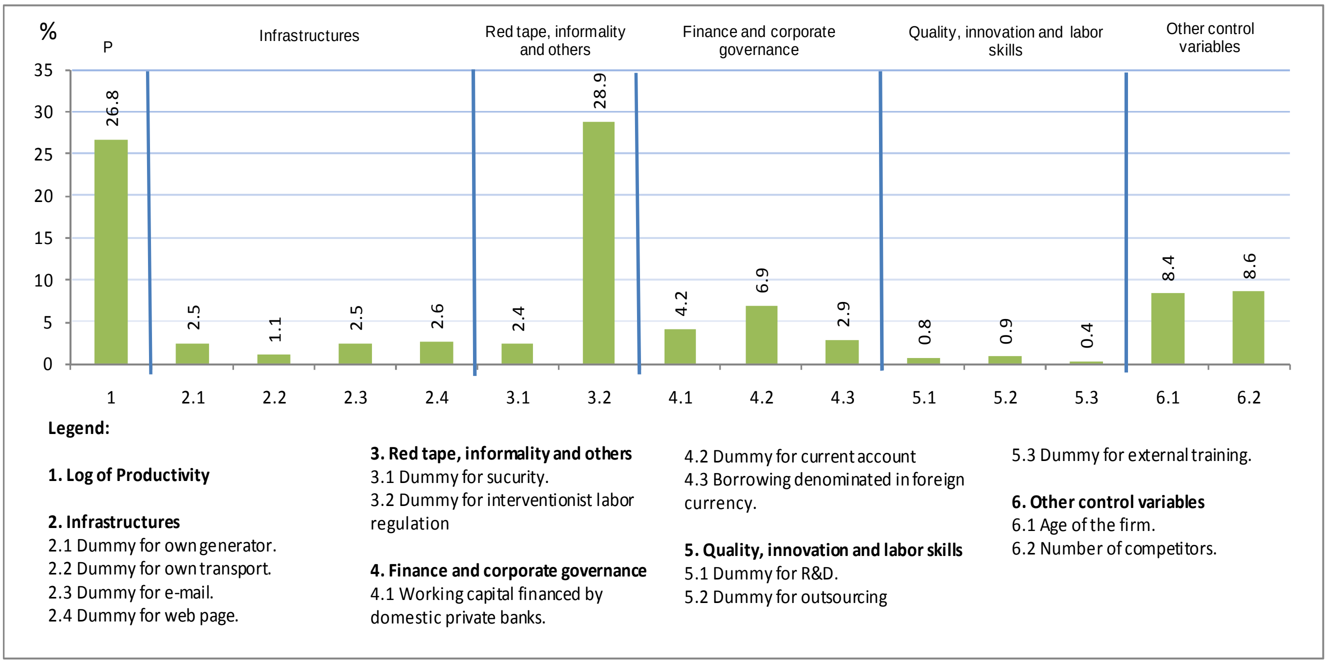
4. Finance and corporate governance
4.1 Largest shareholder.
4.2 New fixed assets financed by domestic commercial banks.
4.3 Dummy for loan.

5. Quality, innovation and labor skills
5.1 Dummy for R&D.
5.2 Staff - skilled workers.
5.3 Dummy for external training.
5.4 Manager tenure.

6. Other control variables
6.1 Dummy for incorporate company.
6.2 Number of competitors.
6.3 Share of imports.

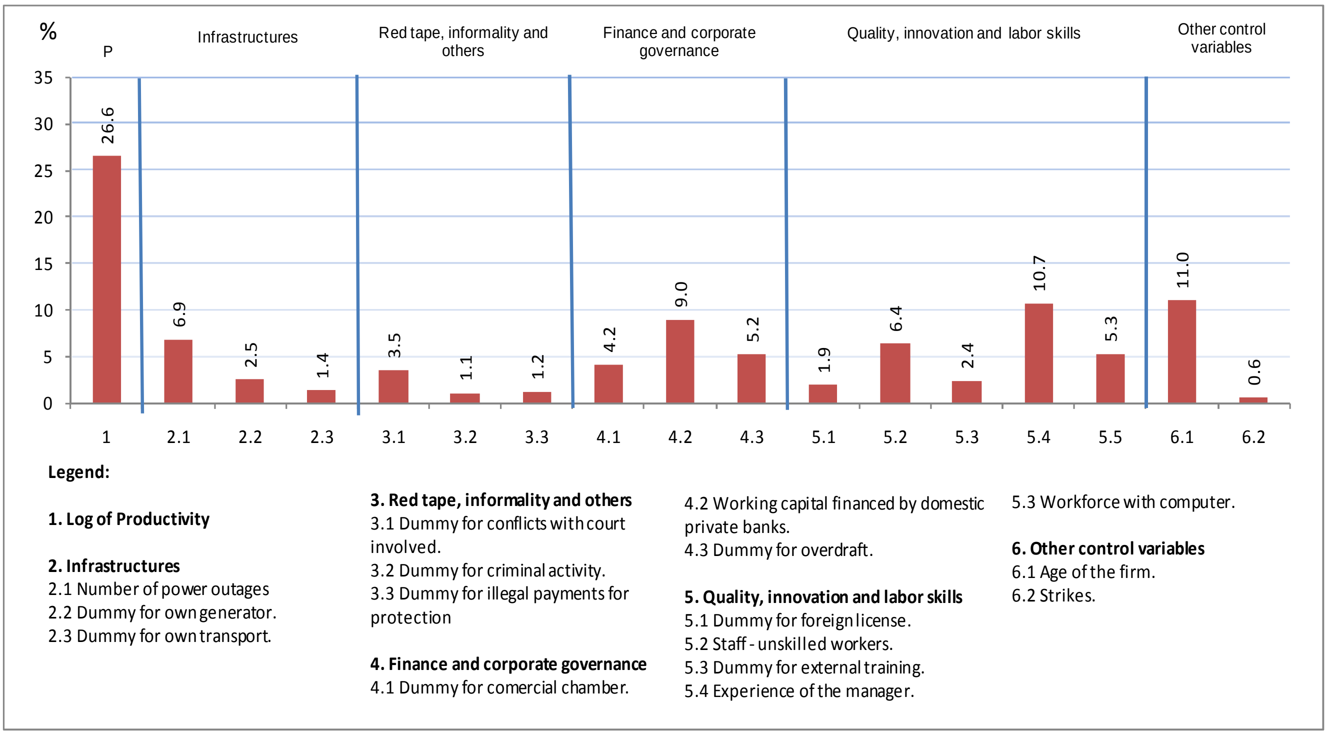
Note: Contributions computed according to section 5.6. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.3: IC percentage contributions to the probability of exporting.



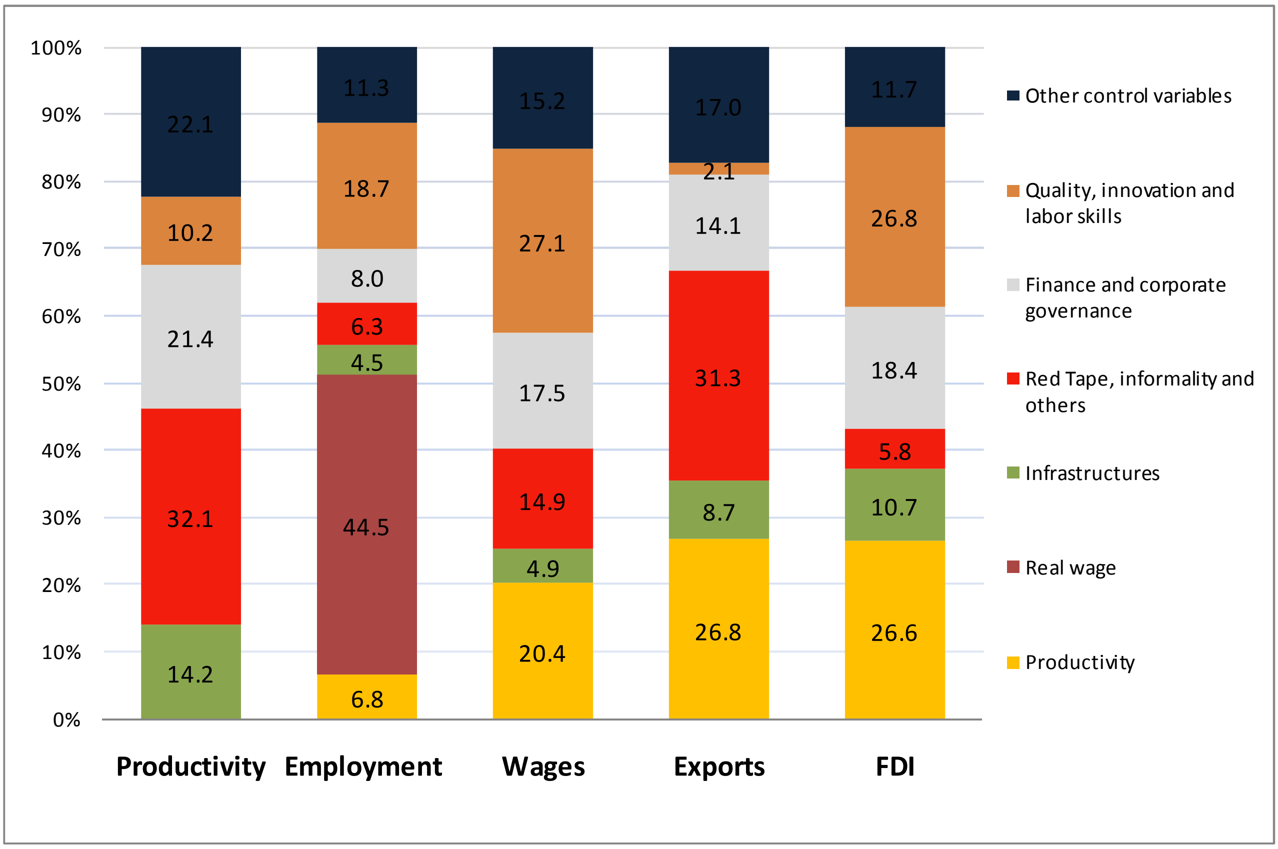
Note: Contributions computed according to section 5.7. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.4: IC percentage contributions to the probability of receiving FDI



Note: Contributions computed according to section 5.8. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

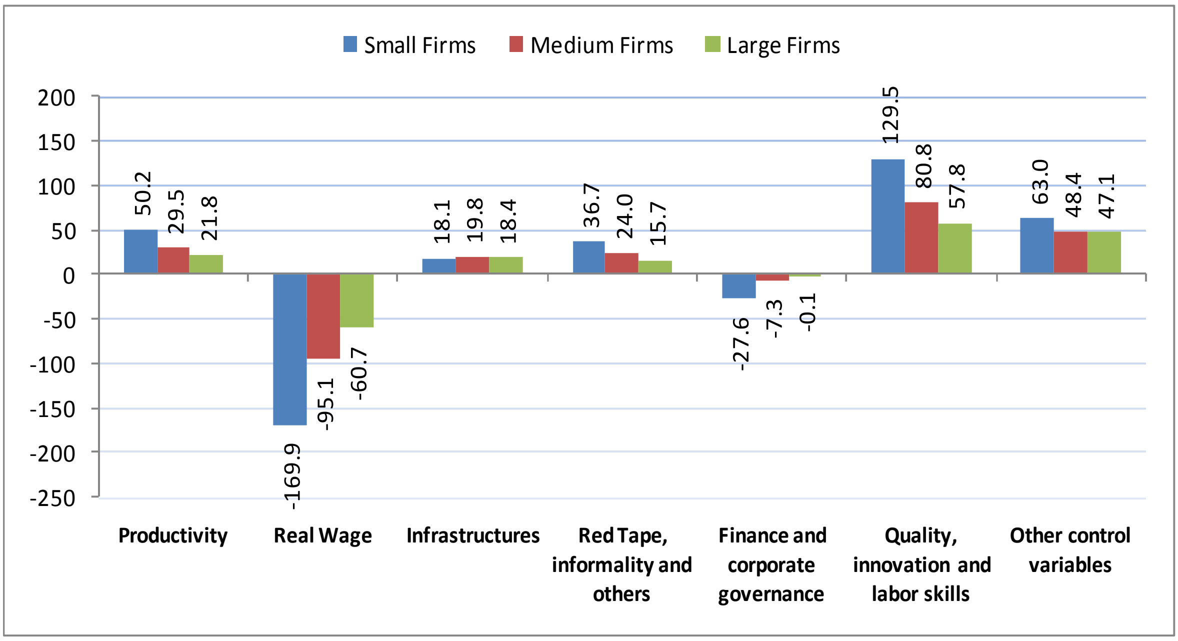
Figure 6: Weight of each block of IC variables on the sample means of economic performance measures



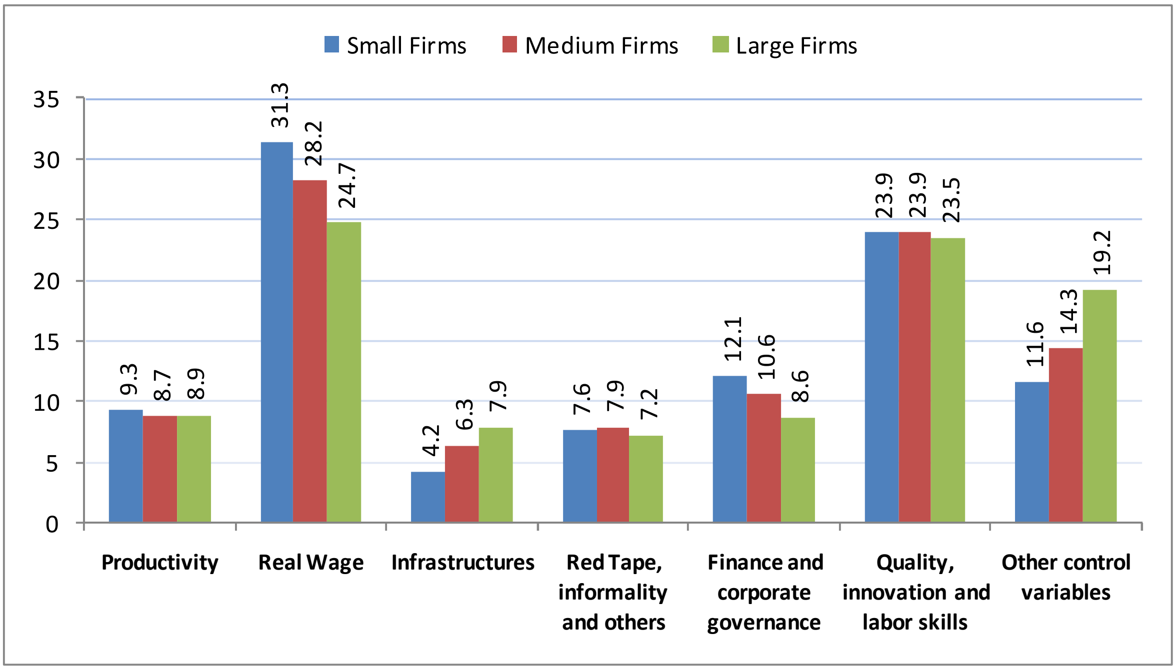
Note: The weight of each block or group of IC variables to the sample means comes from the contributions of Figures comes from figures 5.1 to 5.4. We take the percentage contributions in absolute value and we compute the relative weight of each block. For the case of productivity we take the IC contributions to average log-productivity from Figure 4.3.
Source: Authors' calculations with India ICS data.

Figure 7.1: Weight of each block of IC variables on average log-employment by size.

a) percentage contributions



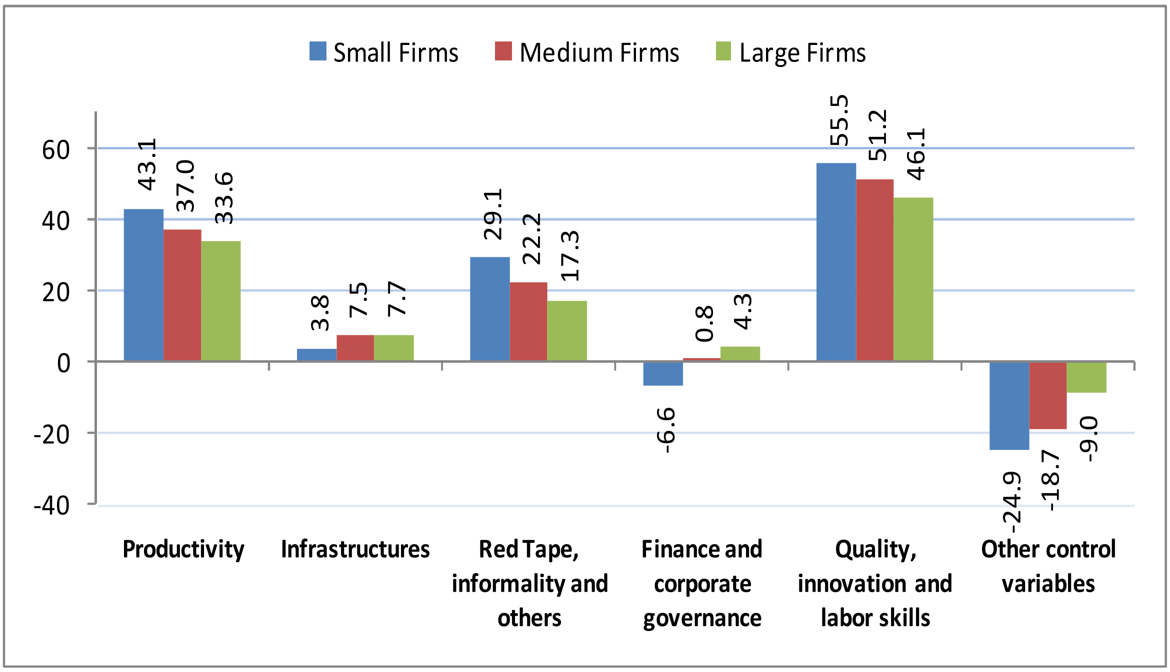
b) absolute percentage contributions



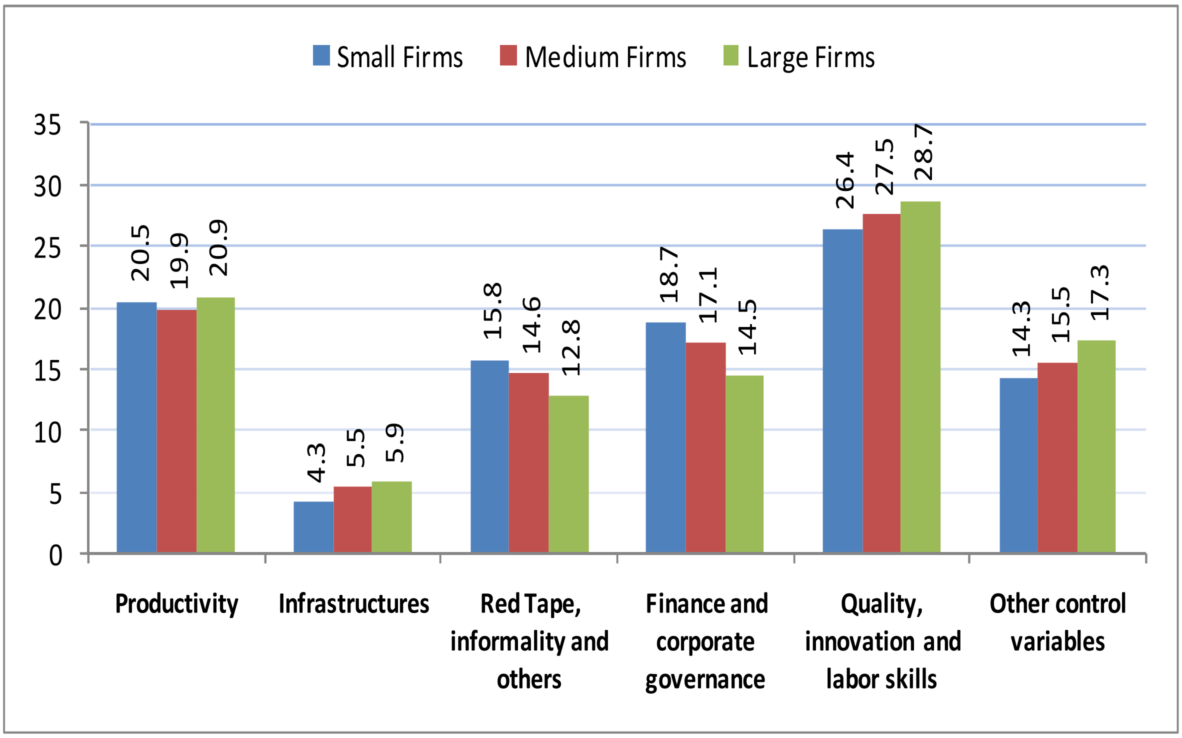
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 7.2: Weight of each block of IC variables on average log-wage by size.

a) percentage contributions



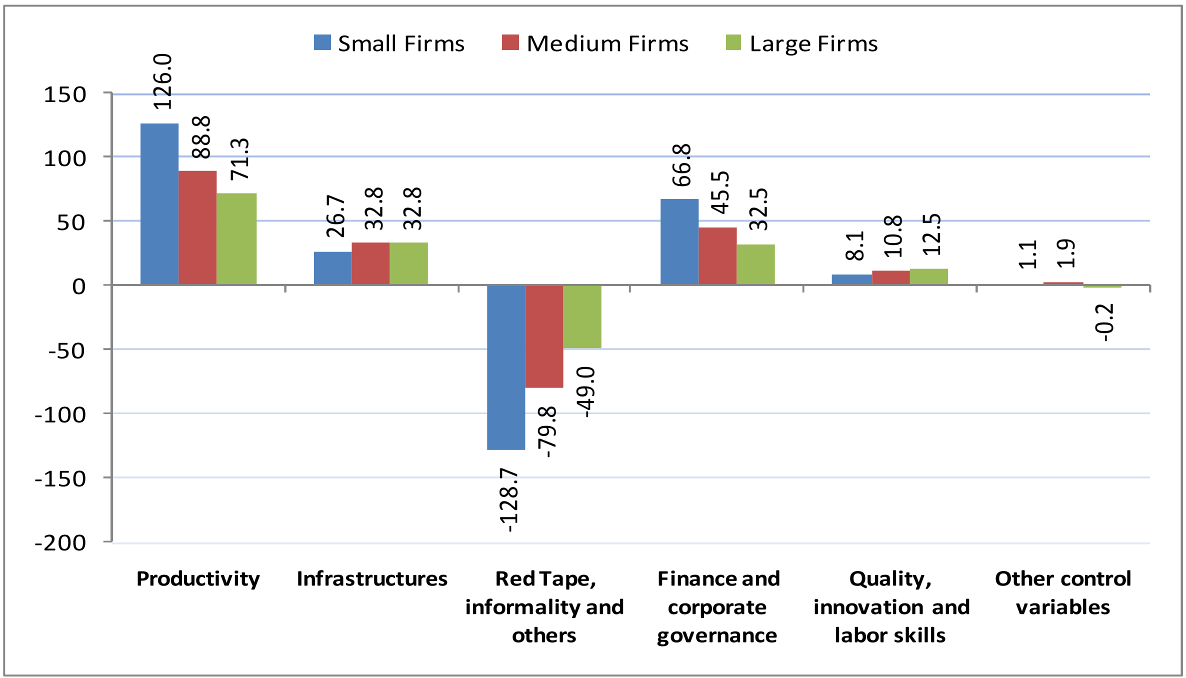
b) absolute percentage contributions



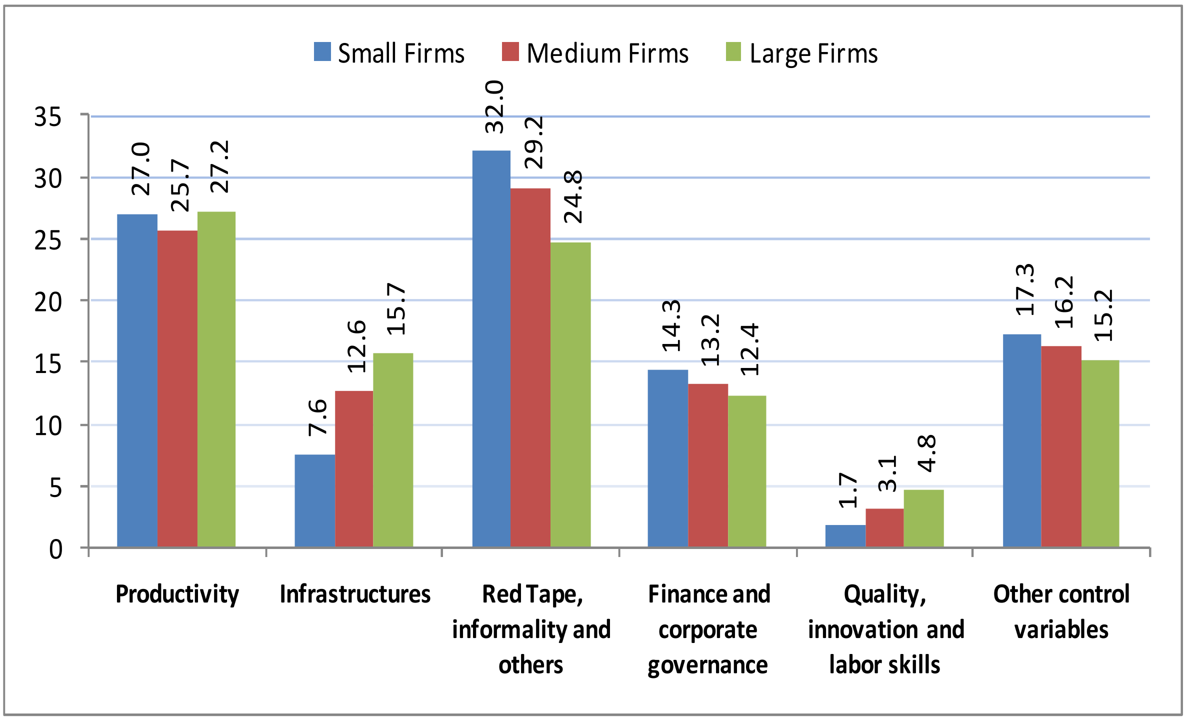
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 7.3: Weight of each block of IC variables on the probability of exporting by size.

a) percentage contributions



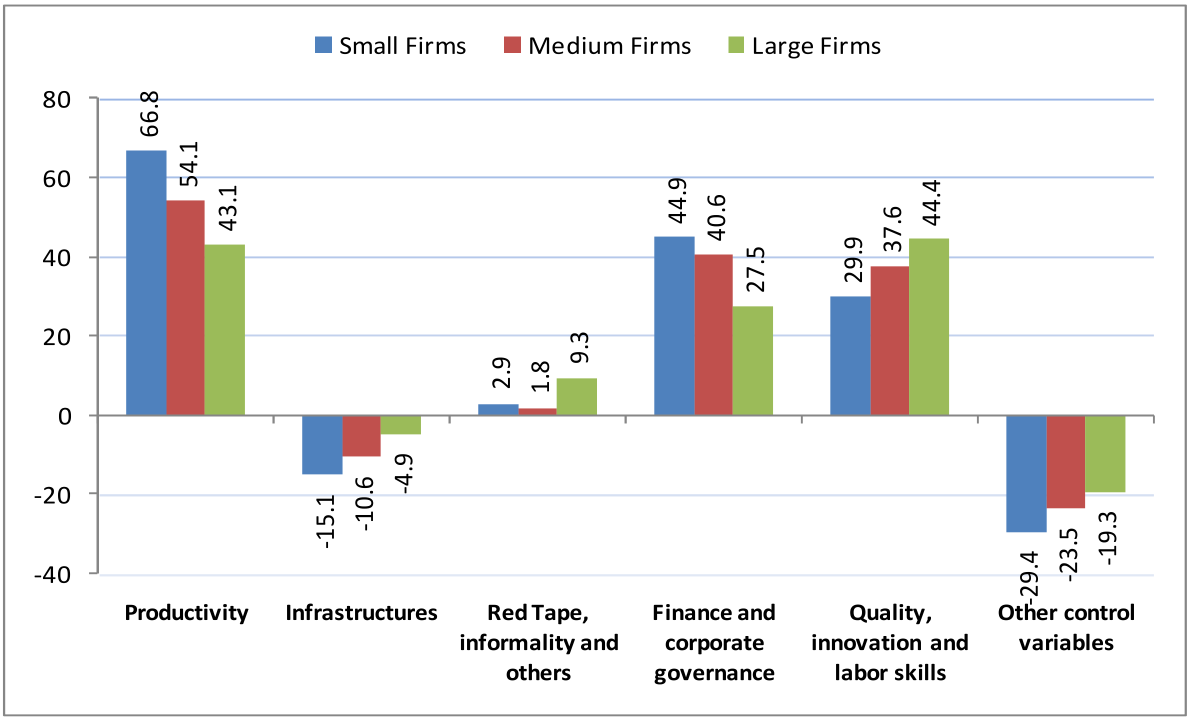
b) absolute percentage contributions



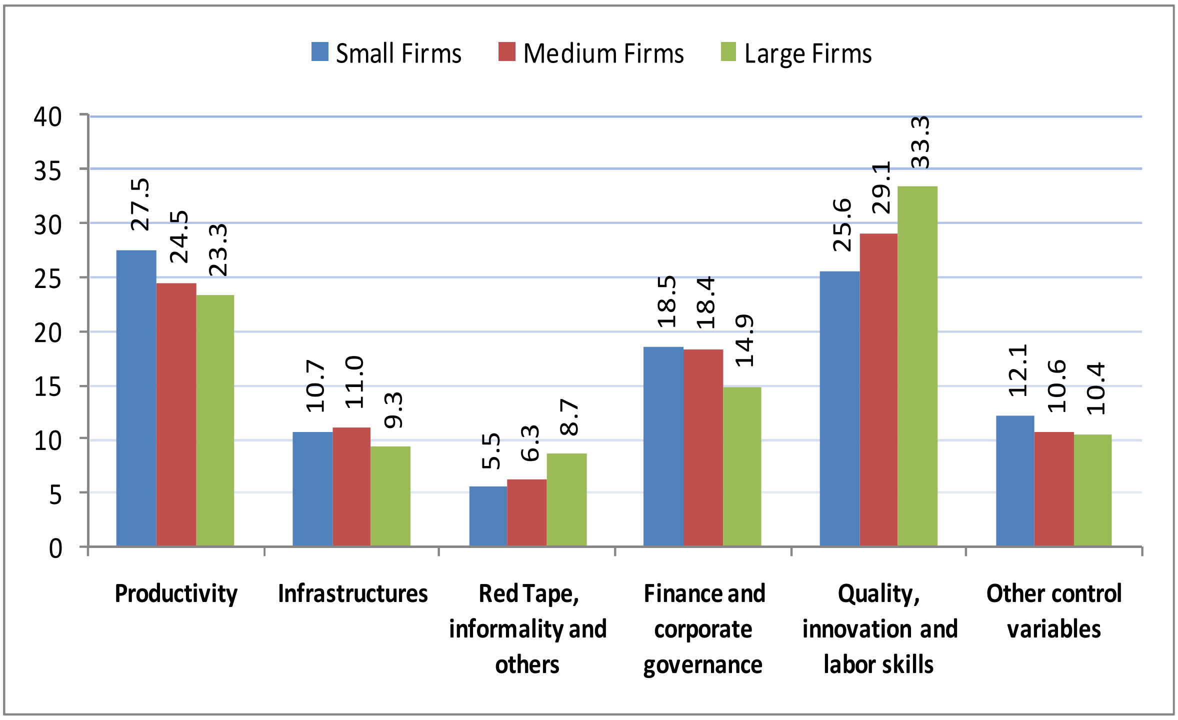
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 7.4: Weight of each block of IC variables on the probability of receiving FDI by size.

a) percentage contributions

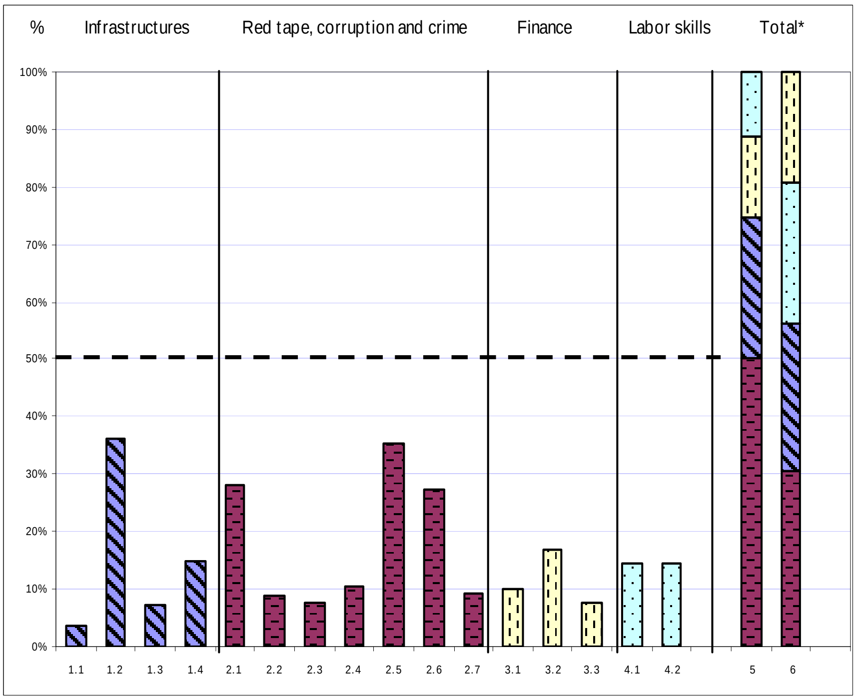


b) absolute percentage contributions



Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 8.1: Firm's perceptions; percentage of firms that considers each one of the following problems as a severe obstacle to firms' economic performance.



1. Infrastructures.
1.1 Telecommunications
1.2 Electricity
1.3 Transportation
1.4 Customs and trade regulations

2 . Red tape, informality and others.
2.1 Corruption
2.2 Crime, theft and disorder
2.3 Anti-competitive or informal practices
2.4 Regulatory policy
2.5 Tax rates
2.6 Tax administration
2.7 Business licensing and operating permits

3. Finance.
3.1 Access to land
3.2 Cost of finance
3.3 Macroeconomic uncertainty

4. Labor skills.
4.1 Labor regulations
4.2 Skills and education of available workforce

5. Total relative weights.
6. Average group relative weights.
* (Totals are computed as the relative weigh of each group of perceptions over the sum of all perceptions' weights)

Source: Authors' calculations with India ICS data.

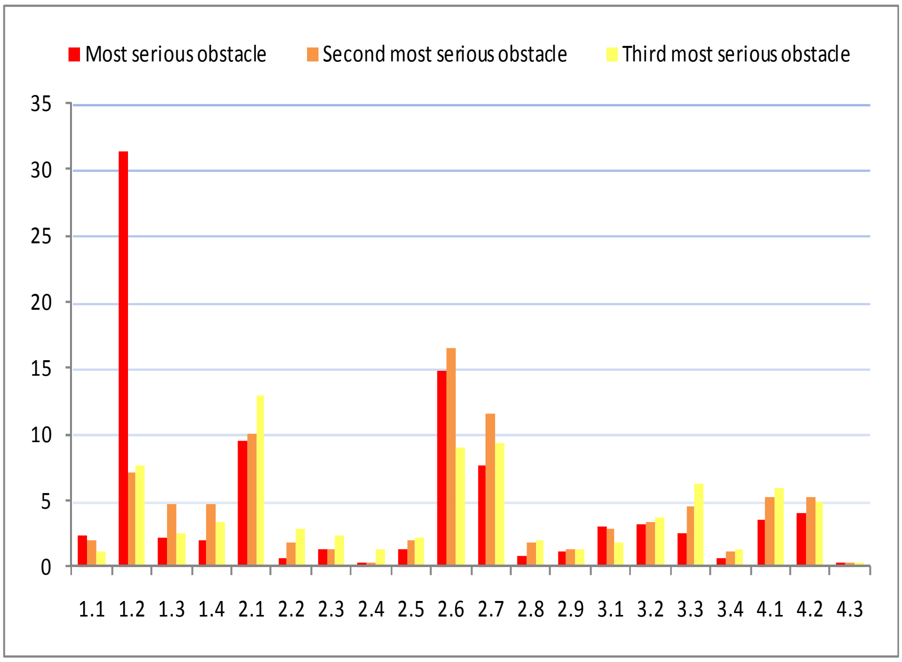
Figure 8.2: Firm's perceptions; first, second and third most serious IC obstacles to firm's economic performance.

1. Infrastructures
1.1 Telecommunications
1.2 Electricity
1.3 Transportation

2. Red tape, informality and others
2.1 Corruption
2.2 Crime, theft and disorder
2.3 Anti-competitive or informal practices
2.4 Legal system/conflict resolution
2.5 Economic and regulatory policy uncertainty
2.6 High rates
2.7 Tax administration
2.8 Business licensing and operating permits
2.9 Regulation specific to your industry

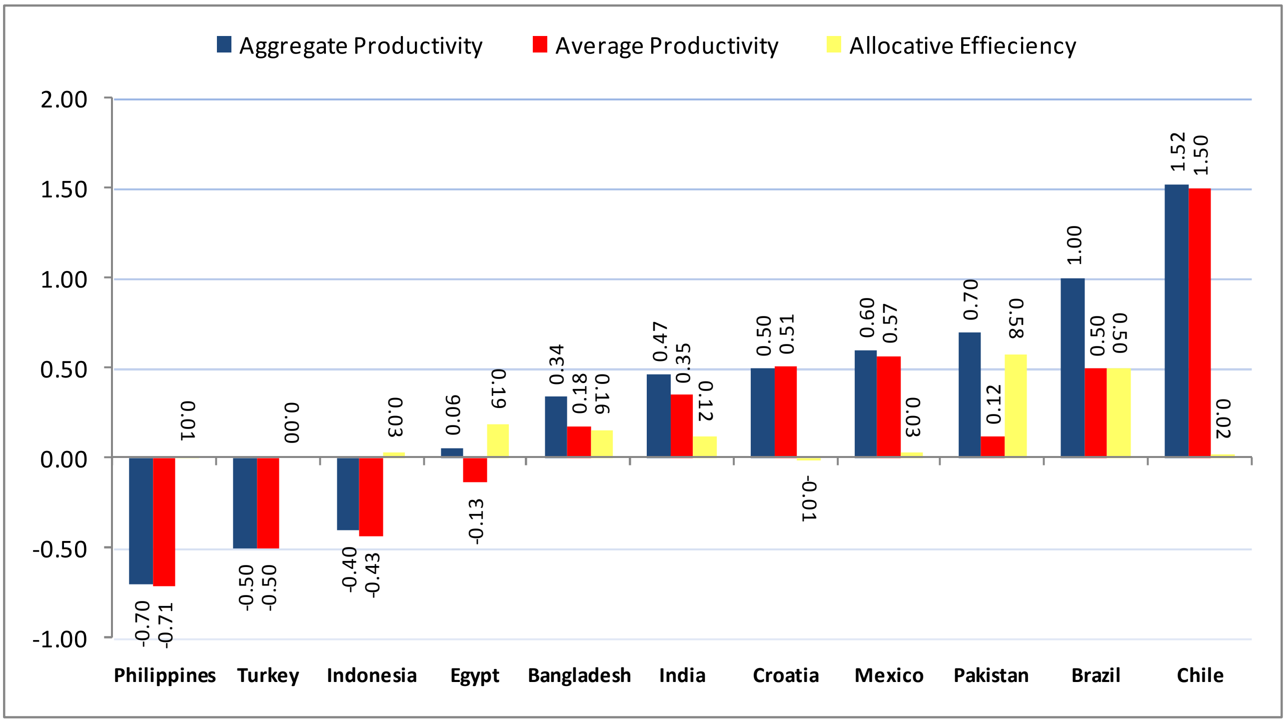
3. Finance
3.1 Access to land
3.2 Access to finance
3.3 Cost of financing
3.4 Macroeconomic uncertainty

4. Labor skills
4.1 Labor regulations
4.2 Skills and education of available workforce
4.3 Access to foreign technology



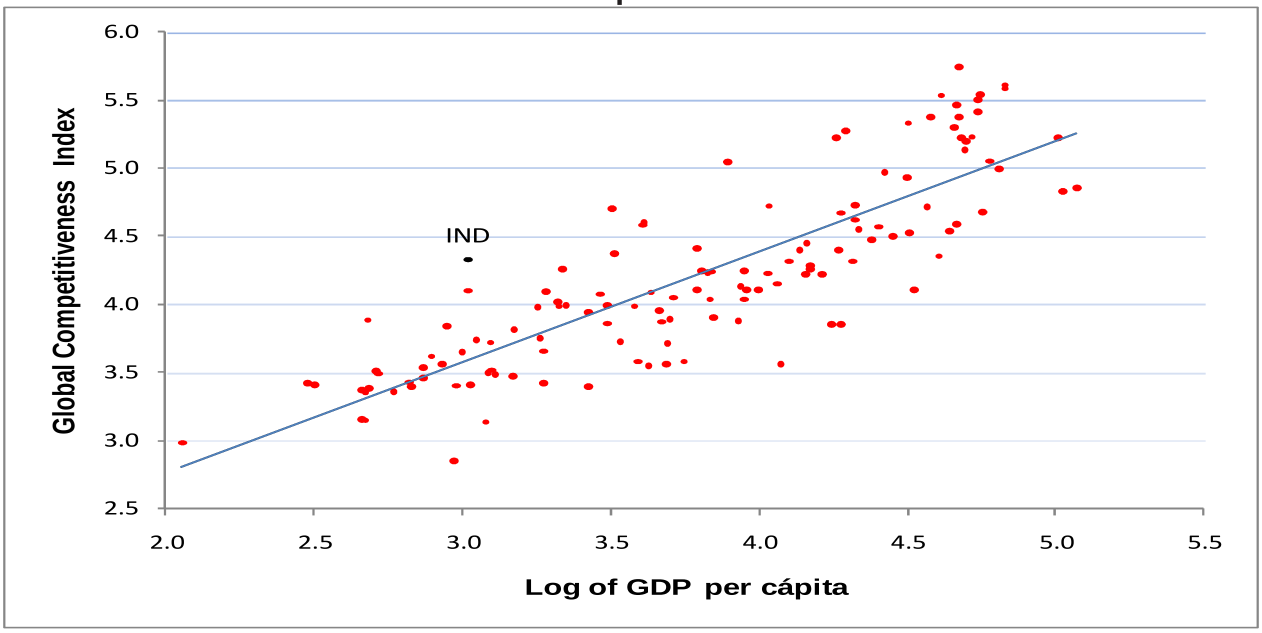
Source: Authors' calculations with India ICS data.

Figure 9.1: *Demeaned mixed* Olley and Pakes decomposition in India and comparators.



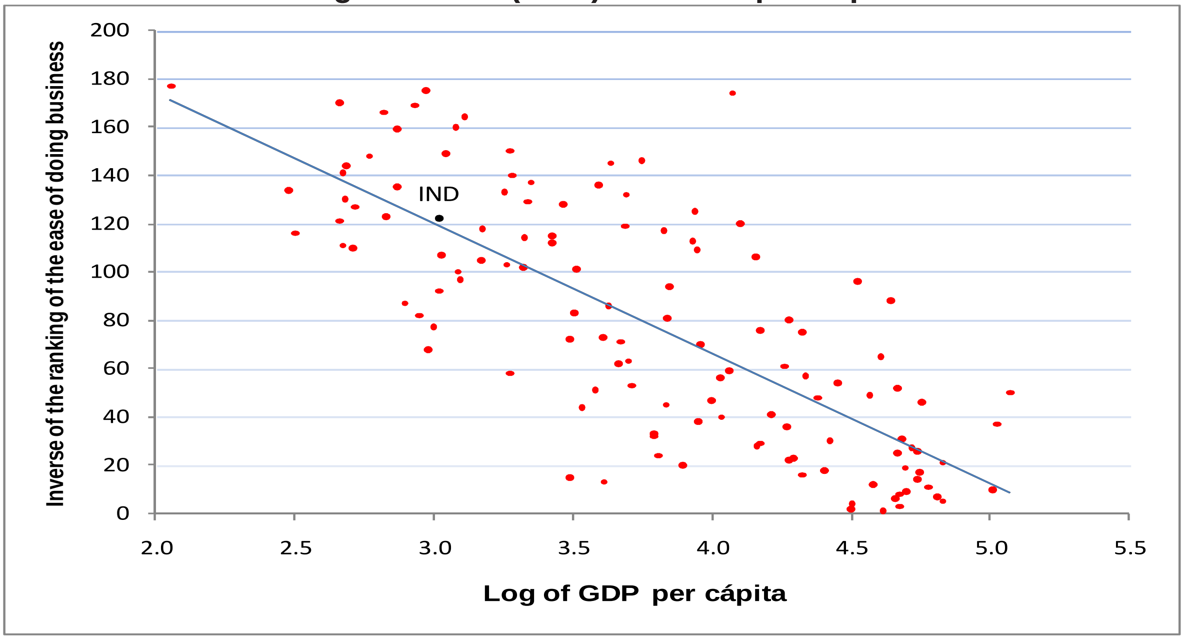
Note: Olley and Pakes decomposition in levels according to equation (5.4).
The productivity measure used is the *demeaned* restricted Solow residual in logs.
Source: Authors' calculations with India ICS data

Figure 9.2: Cross-plot between Global Competitiveness Index (2009) and GDP per capita.



Notes:
The global Competitiveness Index is computed as a weighted average of the 12 fundamental pillars for competitiveness. The stage of development of each economy is taken into account when computing the weights of each pillar.
Source: Authors' elaboration with data from the World Economic Outlook 2009, IMF; y Global Competitiveness Report 2008, The World Economic Forum.

Figure 9.3: Cross-plot between the (inverse) rankings of the ease of doing business (2008) and GDP per capita.



Notes:
The ranking of the ease of doing business is the result of a weighted average of each one of the rankings of the basic aspects of doing business. In the vertical axis it is the inverse of the ranking, say number of countries (181) minus the ranking of each country. The higher the inverse the easier it is doing business.
Source: Authors' elaboration with data from the World Economic Outlook 2009, IMF and Doing Business Report 2009, The World Bank Group, Washington, DC.

Chapter III:

Economic Determinants of India`s Investment Climate

ICT Sector

Economic Determinants of India`s Investment Climate

ICT Sector

Abstract

ICT sector represents the successful of the Liberalization polices of the early 1990`s. The positive results in this sector are based on a combination of market conditions: a favorable domestic investment climate environment and a large demand for Western companies seeking to reduce cost. However, these favorable conditions have allowed productivity in the ICT sector to lag its potential. Based on the methodology for ICs developed by Escribano and Guasch (2005, 2008) and Escribano et al (2008a and 2008b), we identified the main IC constrains affecting productivity and other economic performance measures of ICT firms in India. Our econometric analysis using The ICT ICs 2006 (From the World Bank) highlights key investment climate constrains in the low quality of the supply of power, bureaucracy and corruption and in the shortage of qualified workforce.

JEL Clasification: D24; D61; J20; J30; L86; O40; O53

Keywords: *ICT, productivity, investment climate, exporter, average contribution.*

Table of contents

1. INTRODUCTION 105

2. DATA 106

3. ECONOMETRIC ESTIMATION OF IC ELASTICITIES AND SEMI-ELASTICITIES ON PRODUCTIVITY (TFP). 107

3.1 Robustness of IC elasticities and semi-elasticities: single step and two step estimation, restricted and unrestricted input-output elasticities. 108

3.2 Endogeneity of production function (PF) variables. 110

3.3 Role of prices on production function (sales generating functions and market power). 110

3.4 Endogeneity of IC variables. 111

3.5 Selection of the relevant models. 113

4. ECONOMETRIC ANALYSIS OF IC AND PRODUCTIVITY IMPACT ON EMPLOYMENT, REAL WAGES, PROBABILITY OF EXPORTING AND PROBABILITY OF RECEIVING FDI. 113

4.1 Identification of the system of equations 116

5. IC ASSESSMENT ON AGGREGATE PRODUCTIVITY (OLLEY AND PAKES DECOMPOSITION) AND OTHER MEASURES OF ECONOMIC PERFORMANCE. 117

5.1 O&P decompositions: in levels and mixed. 118

5.2 IC effects on productivity measure in the terms of the mixed O&P decomposition. 118

5.3 Simulations based on the IC effects on the O&P decomposition of TFP. 119

5.4 IC evaluation on the sample means of employment and wages, on the probability of exporting and on the probability of receiving FDI. 120

6. EMPIRICAL RESULTS. 121

6.1 Key Results on Productivity 121

Exporters versus non-exporters 123

6.2 Key Results on Employment 123

6.3 Key Results on Real Wages 125

6.4 Key Results on the Probability of Exporting 126

6.5 Key Results on the Probability of Receiving Foreign Direct Investment 127

7. CONCLUSIONS. 128

REFERENCES 130

APPENDIX A: DEFINITIONS 133

APPENDIX B: TABLES AND FIGURES 138

1. Introduction

As developing countries face the pressures and impacts of globalization, they are seeking ways to stimulate growth and employment within this context of increased openness. With most of these countries having secured a reasonable level of macroeconomic stability, they are now focusing on issues of competitiveness and productivity through microeconomic reform programs. Governments are reformulating their strategies and making increased competitiveness a key priority of government programs.

A main component of country competitiveness is having a good investment climate or business environment. The investment climate, as defined in the WDR (2005), is “the set of location-specific factors shaping the opportunities and incentives for firms to invest productively, create jobs and expand.”

Government policies and behavior exert a strong influence on the investment climate through their impact on costs, risks and barriers to competition. Key factors affecting the investment climate through their impact on costs are: corruption, taxes, the regulatory burden and extent of red tape in general, factor markets (labor, intermediate materials and capital), the quality of infrastructure, technological and innovation support, and the availability and cost of finance.

While the Investment Climate Assessments are quite useful in identifying major issues and bottlenecks as perceived by firms, the data collected is also meant to provide the basic information for an econometric assessment of the impact or contribution of the investment climate (IC) variables on productivity.

We believe that improving the investment climate (IC) is a key policy instrument to promote economic growth and to mitigate the institutional, legal, economic and social factors that are constraining the convergence of per capita income and labor productivity of India relative to more developed countries. For that, we need to identify the main investment climate variables that affect economic performance measures like total factor productivity, employment, wages, exports and foreign direct investment and this is the main goal of this paper.

In 1998, OECD member countries agreed to define the ICT sector as a combination of manufacturing and services industries that capture, transmit and display data and information electronically.

For the study of the economy of India, we have to take in to account the diversity of the country, multifaceted nation with a variety of topographies, climates, cultures, languages and work ethics.

Work environments and cultures in various parts of the country can be as diverse as in the different countries of Europe.

ICT has been the fastest growing sector and is the largest contributor to exports. Within this sector, India is especially successful in software sub-sector, being one of the world superpower as Gartner Report 2006 illustrates. ICT sector represents the successful of the Liberalization polices of the early 1990`s. The positive results and achievements in this sector are based on a combination of market conditions: a favorable domestic investment climate environment and a large demand for Western companies seeking to reduce cost.

However, these favorable conditions have allowed productivity in the ICT sector to lag its potential. Identify key investment climate bottlenecks that reduce the potential growth of ICT sector in India will be useful to achieve the future goals.

The core of the methodology is a structural system of equations relating investment climate as right hand side (explanatory) variables and productivity, employment, wages, probability of exporting and probability of receiving FDI as endogenous or left hand side variables. In the methodological aspects of the estimation of the system we follow Escribano and Guasch (2005 and 2008) and Escribano et al (2008b). Once we have identified the significant IC effects on economic performance we evaluate the IC productivity contributions in terms of the Olley and Pakes (1996) decomposition of aggregate productivity and the IC contributions in terms of the sample means of the remaining economic performance measures (see Escribano et al. 2008a and Escribano, Guasch and Pena, 2008).

The results of the econometric analysis show a clear relation between IC and economic performance in India. We find that infrastructure factors like the quality of power supply and informalities are significantly associated with the large differences in productivity observed in the sample of establishments used. In addition, we observe that exporter firms have in the limited skills workforce other important obstacle while for non-exporter firms have its specific bottleneck in the access to land.

2. Data

The World Bank Group in close partnership with public or private institutions in each country creates fir level data for investment climate assessments. The surveys of private enterprises try to find out about the difficulties that firms encounter in starting and running a business—and, if the business fails, in exiting. The ICs survey captures firms' experience in a range of areas—financing, governance, regulation, tax policy, labour relations, conflict resolution, infrastructure services, supplies and marketing, technology, and training.

From India´s IC survey we are able to form a cross section data base. The cross section is for the fiscal year 2004-2005. The survey covered 359 firms in the ICT sector. The firms were located mainly in the seven largest cities, which had the highest concentration of ICT firms.

3. Econometric estimation of IC elasticities and semi-elasticities on productivity (TFP).

In the identification of the significant investment climate effects on economic performance (productivity, demand for labor, real wages, probability of exporting and probability of receiving FDI) it is important to condition on the whole set of information contained in the IC survey. In particular, we propose a simultaneous equations system that relates the interactions between the investment climate variables and firm's economic performance measures.

Escribano and Guasch (2005, 2008), model that relates IC and C variables with firm-level productivity (TFP) **by the following system of equations with fixed effects**,

$$\log Y_{it} = \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \log TFP_{it} \qquad (3.1a)$$

$$\log TFP_{it} = a_i + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + w_{it} \qquad (3.1b)$$

$$a_i = \alpha'_{IC} IC_{P,i} + \alpha'_C C_{P,i} + \varepsilon_i \qquad (3.1c)$$

where, *Y* is firms' output (sales), *L* is employment, *M* denotes intermediate materials, *K* is the capital stock, *IC* and *C* are time-fixed effect vectors of other investment climate and control time-fixed effects, and D_r, D_j and D_t are the vectors of state, industry and year dummies.

The usually unobserved time fixed effects (a_i) of the TFP equation (3.1b) are here proxy by the set of observed time fixed components IC, and C variables of (3.1c) and a remaining unobserved random effects (ε_i). The two random error terms of the system, ε_i and w_{it}, are assumed to be conditionally uncorrelated with the explanatory L, M, K, IC and C variables[16] of equation (3.2),

[16] Under this formulation (and other standard conditions) the OLS estimator of the productivity equation (3.2) with robust standard errors is consistent, although a more efficient estimator (GLS) is given by the random effects (RE) estimator that takes into consideration the

$$\log Y_{it} = \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \alpha'_{IC} IC_{P,i} + \alpha'_C C_{P,i} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + u_{it} \qquad (3.2)$$

Therefore, the regression equation (3.2) represents the *conditional expectation* plus a composite random-effect error term equal to $u_{it} = \varepsilon_i + w_{it}$.

Before introducing the remaining equations of the system we explain the main econometric issues that we have to address in the estimation of productivity (TFP) equations.

3.1 Robustness of IC elasticities and semi-elasticities: single step and two step estimation, restricted and unrestricted input-output elasticities.

By simply plugging (2.1c) into (2.1b) we get the next expression for productivity

$$\log TFP_{it} = \alpha'_{IC} IC_i + \alpha'_C C_i + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + u_{it} \qquad (3.3)$$

where *IC* and *C* are, respectively, the observable fixed effects vectors of investment climate and control variables listed in Tables 2.1 to 2.7 of the Appendix. In the regressions, we always control for several region dummies (D_r, r= 1, 2… R), sector-industry dummies (D_j, j = 1, 2... q_D), a constant term (α_P) and in the panel data case we also include a set of time dummies (D_t, t = 1, 2... q_T). Since there is no single salient measure of productivity (or $logTFP_{it}$), any empirical evaluation of the productivity impact the IC might critically depend on the particular productivity measure used. Escribano and Guasch (2005, 2008) suggested–following the literature on *sensitivity analysis* of Magnus and Vasnev (2006)–to look for empirical results (elasticities) that are *robust to several productivity measures*. This is also the approach we follow in this paper.

In particular, we want the elasticities of IC on productivity (TFP) to be robust (with equal signs and similar magnitudes) for the 6 different productivity measures used. The alternative productivity measures used come from considering:

a) different functional forms of the production functions (Cobb-Douglas and Translog),

b) different sets of assumptions (technology and market conditions) to get consistent estimators based on Solow's residuals by ordinary least squares (OLS).

---

particular covariance structure of the error term, $\varepsilon_i + w_{it}$, which introduces a particular type of heteroskedasticity in the regression errors of (3.2).

c) different aggregation levels when measuring input-output elasticities (industry level or aggregate country level).

Table A: Summary of productivity (P) measures and estimated investment climate (IC) elasticities

Functional forms of production function	Estimation procedure	Aggregation level of coefficients of PF	Result
1. Solow´s Residual	Two-step estimation	1.1 Restricted coefficients	2 (TFP) measures; 2 (IC) elasticities
		2.2 Unrestricted coefficients	
2. Cobb-Douglas	Single-step estimation	2.1 Restricted coefficients	2 (TFP) measures; 2 (IC) elasticities
		2.2 Unrestricted coefficients	
3. Translog	Single-step estimation	3.1 Restricted coefficients	2 (TFP) measures; 2 (IC) elasticities
		3.2 Unrestricted coefficients	
Total			**6 (TFP) measures and therefore 6 estimates of IC elasticities (or semi-elasticities)**

Note: Restricted coefficient = equal input-output elasticities in all industries.
Unrestricted coefficient = different input output elasticities by industry.

Table A above summarizes the productivity measures used for the IC robust evaluation. The two-step estimation starts from the nonparametric approach based on cost shares from Hall (1990) to obtain Solow's residuals in logs under two different assumptions:[17] (a) the cost shares are constant for all plants located in the same country (restricted Solow residual), and (b) the cost shares vary among industries in the same country (unrestricted by industry Solow residual). Once we have estimated the Solow residuals ($logTFP_{it}$) in the first step, in the second step we can estimate equation (3.3) by OLS with robust standard errors and allowing for clustering correlation within industries and states.

In the single-step estimation approach, we start with the OLS parametric estimation of the extended production function (3.2). We use two different functional forms of the PF—Cobb-Douglas and Translog—under two different aggregation conditions on the input-output elasticities: equal input-output elasticities in all industries (restricted case) and different input-output elasticities by industries (unrestricted case).

[17] The advantage of the Solow residuals is that they require neither the inputs (L, M, K) to be exogenous nor the input-output elasticities to be constant or homogeneous (Escribano and Guasch, 2005 and 2008). The drawback is that they require having constant returns to scale (CRS) and, at least, competitive input markets.

3.2 Endogeneity of production function (PF) variables.

There is an identification issue separating TFP from PF when any PF inputs is influenced by unobserved common causes affecting productivity—such as a firm's fixed effects. This creates simultaneous equation bias if least squares are used estimating equation (3.1a) to measure TFP. However, this endogeneity problem of the inputs is overcome by using the single step least squares estimation of equation (3.2) follow the approach proposed by Escribano and Guasch (2005, 2008). That is, in (3.2) we proxy the usually unobserved firm-specific fixed effects (which are the main cause of inputs' endogeneity) by a long list of observed firm-specific fixed effects coming from the investment climate surveys. Controlling for the largest set of IC variables and plant C characteristics, we can—under standard regularity conditions— get consistent and unbiased least squares estimators of the parameters of the PF and the corresponding IC elasticities on TFP in one step.

Notice that even if we were only interested in assessing the impact of one block of IC variables, say infrastructure, we do not limit the scope of the analysis to only that block of IC variables. We include (and therefore control for) IC factors from all the blocks because of the crucial role IC variables play as proxies for the unobserved fixed effects. This is the key feature of the Escribano and Guasch (2005, 2008) econometric methodology to provide robust empirical regularities. If for example, we try to estimate the impact of say infrastructure, without controlling for the other IC blocks of variables, we can get different signs on certain coefficients due to the omitted variables problem; see Escribano and Guasch (2008).

3.3 Role of prices on production function (sales generating functions and market power).

The role of prices in the system (3.1a)-(3.1c) deserves special attention. As our dependent variable is sales, rather than units of physical output, it reflects prices. In fact, according to the current literature, the term *sales generating function* seems more appropriate rather than production function for equation (3.1a), as in the work of Olley and Pakes (1996). If prices are not identical across firms, what seems to be a high productive plant may be just an establishment that is charging high prices, what in turn may be consequence of either market power (non zero mark-ups) or differences in quality of final goods. While with homogeneous products high productivity could be a reflection of high prices, or in other words a reflection of market power (Melitz, 2000; Bernard, et al., 2003; Katayama, et al., 2006; Foster et al, 2008), under heterogeneous or differentiated products high prices could be consequence of higher quality, what could be translated to over-measured productivity as some plants would be able to produce higher quality—and price—products with the same amount of output (Levinsohn and Melitz, 2002; de Loecker, 2007; Katayama, et al., 2006; Gorodnichenko, 2007). These points are especially important in developing countries where usually market power is a severe constraint to growth. Addressing these issues is not a straightforward task with the data available. A more comprehensive analysis would need information on plant level input prices to incorporate the demand side of the model.

As long as this data is not available a plausible solution is to estimate the system (3.1a)-(3.1c) by following a control approach. Now instead of observing output (Y) we are observing sales (P_yY), where P_y denotes prices, and then equation (3.1a) is transformed to (3.1a')

$$\log Y_{it} + \log P_{y,it} = \log P_{y,it} + \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \log TFP_{it} \quad (3.1a')$$

Notice that as long as we control for *logPy* on the right hand side of equation (3.1a'), *productivity* in the RHS of the equation still is *logTFP*. Since, within a year there is low price variability at the firm level we assume that *log*P_y can be proxied by a constant term, control variables *C* that are time-firm level fixed effect vectors of firm variables and a set of dummy variables, and D_r, D_j and D_t including the vectors of state, industry and year dummies. Therefore, after including all those variables we could assume that that $\log P_{y,it} \approx \alpha'_C C_{P,i} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t$ and therefore we can get a similar expression for (3.2) incorporating prices

$$\log Y_{it} + \log P_{y,it} = \alpha_L \log L_{it} + \alpha_M \log M_{it} + \alpha_K \log K_{it} + \alpha'_{IC} IC_{P,i} + \alpha'_C C_{P,i} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + u_{it} \quad (3.2')$$

Estimating sales in (3.2'), as we do in our empirical analysis, can provide evidence that TFP can be "interpreted" as "technical efficiency".[18] Finally, to control for the mark up (market power effect) and/or quality (differentiated products) we are also including several IC and C variables related to competition (see the list of IC variables included in the group of other control variables).

3.4 Endogeneity of IC variables.

Another econometric problem we have to face when estimating the parameters of IC, and C variables—either from the two-step or single-step procedure—is the possible endogeneity of some of these explanatory variables. That is, many IC variables are likely to be determined simultaneously along with any TFP measure. With these productivity equations, the traditional instrumental variable (IV) approach is difficult to implement, given that we only have information for one year, and therefore we cannot use natural instruments, such as those provided by their own lags. As an alternative correction for the endogeneity of the IC variables, we use the region-industry-size

[18] Notice, however, that the word technical efficiency that you use is too narrow in the ICs context since there are many efficiencies related to IC variables on TFP that are not technical (regulatory, governance, institutional, etc.).

average of plant-level IC variables instead of the crude IC variables,[19] which is a common solution in panel data studies at the firm level[20].

However, one should avoid including too many industry-region-size variables since it may lead to multicollinariety problems. Especially, if the number of states, sizes and industries is not large enough and there are common regions and/or industries processes affecting the variables. So a proper *a priori* consideration of the endogeneity of IC and C variables is important.

Using industry-region-size averages also mitigates the effect of having certain missing individual IC observations at the plant level, which—as mentioned in Section 2—represent one of the most important difficulties using ICSs. As an alternative, we also follow a second strategy to deal with the missing values of some IC, and C variables. In order to keep as many observations in the regressions as possible to avoid losing efficiency, when the response rate of the variables is large enough, we decided to replace those missing observations with the corresponding industry-region-size average.[21] Thus, we gain observations, efficiency, and representativity maybe at the cost of introducing some measurement errors into the explanatory variables.[22]

For those variables which endogeneity is intrinsic due to the construction of the simultaneous system of equations (exporting probability and probability of receiving FDI inflows) we apply standard IV estimators (2SLS) using as instruments either the industry-region-size average or those exogenous IC variables from the list of explanatory covariates of the corresponding equation.

Unfortunately, endogeneity is yet an unsettled issue in ICSs. Implementation of those techniques that allow obtaining causal interpretations, like those derived from the concept of '*Granger causality*' or experimental or quasi-experimental methods, are unfeasible to implement in the actual context of IC surveys with cross-sectional dataset or with incomplete panels with a very short time dimension. Although the solutions proposed to deal with endogeneity in this report can reduce the degree of endogeneity of both IC and PF variables, they do not allow us to place causal interpretations on the results obtained. Rather, we have to satisfy ourselves by obtaining empirical regularities with the relationships among IC variables and measures of firms' economic performance.

[19] For the creation of cells a minimum number of firms are imposed—there must be at least 15 to 20 firms in each industry-region-size cell to create the average, otherwise we apply the region-industry averages. If the problem persists, we apply the industry-size or the region-size average.
[20] This two-step estimation approach is a simplified version of an instrumental variable estimator (two-stage least squares, 2SLS).
[21] Notice that this replacement strategy has a straightforward weighted least squares interpretation since we are giving a greater weight to those observations with more variance (Escribano at al., 2008b).
[22] Depending on the assumption we make, the measurement error may introduce a downward bias in the parameters that depends on the ratio between the variances of the variables and the measurement error. Since those explanatory variables are constant within regions, sizes, and industries we expect their variances will be small.

3.5 Selection of the relevant models.

The econometric methodology applied for the selection of the variables (IC, and C) goes from the general to the specific. The otherwise omitted variables problem that we encounter—starting from a too-simple model—generates biased and inconsistent parameter estimates. We start the selection of IC variables with a wide set compounded by up to 90 variables. We avoid using at the same time in the regression, explanatory IC variables that provide similar information (highly correlated), mitigating the problem of multicollinearity that could otherwise arise. We then start removing from the regressions—the less significant variables—one by one, until we obtain the final set of IC variables, significant in at least one of the alternative TFP regressions and with parameters varying within a reasonable range of values. Once we have selected a preliminary model we test for omitted IC variables (those initially dropped IC variables).

The robust TFP effects obtained on IC and C variables, along with their level of significance, are listed in Table C.I of the appendix included at the end of the report. Indications of the form the variables are entering the regression—industry-region-size average or missing values replaced by the industry-region-size average, logs, etc.—are also included in the Table. In all the cases we are using robust standard errors.

4. Econometric analysis of IC and productivity impact on employment, real wages, probability of exporting and probability of receiving FDI.

The same idea of approximating the unobservable fixed effect by the firm level investment climate conditions is applied in the remaining equations of the model.

The demand for labor determined by firm level productivity (*$logP_{it}$*) and by real wages in logs (*$logW_{it}$*) and is given by;

$$log L_{it} = \gamma_L + a_{L,i} + \gamma_P \log TFP_{it} + \gamma_w \log W_{it} + \gamma_{Exp} y_{it}^{Exp} + \gamma_{FDI} y_{it}^{FDI} + \gamma'_{DR} D_r + \gamma'_{Ds} D_j + \gamma'_{DM} D_m + \gamma'_{DT} D_t + \varepsilon_{L,it} \tag{4.1a}$$

$$a_{L,i} = \gamma'_L IC_i^L + \gamma'_C C_i^L + v_{L,i} \, . \tag{4.1b}$$

The wage equation is determined by the productivity (TFP) level after controlling for all the IC effects and by the fact that certain firms exports and receive FDI;

$$logW_{it} = \beta_W + a_{W,i} + \beta_P \log TFP_{it} + \beta_{Exp} y_{it}^{Exp} + \beta_{FDI} y_{it}^{FDI} + \beta'_{DR} D_r + \beta'_{Ds} D_j + \beta'_{DM} D_m + \beta'_{DT} D_t + \varepsilon_{W,it} \quad (4.2a)$$

$$a_{W,i} = \beta'_{IC} IC_i^W + \beta'_C C_i^W + v_{W,i}\,. \quad (4.2b)$$

The probability of firms entering the export market depends on firm level productivity (TFP), the investment climate and by the fact that certain firms receive FDI;

$$y_{it}^{Exp} = \delta_{Exp} + a_{Exp,i} + \delta_P \log TFP_{it} + \delta_{FDI} y_{it}^{FDI} \delta'_{DR} D_r + \delta'_{Ds} D_j + \delta'_{DM} D_m + \delta'_{DT} D_t + \varepsilon_{Exp,it} \quad (4.3a)$$

$$a_{Exp,i} = \delta'_{IC} IC_i^{Exp} + \delta'_C C_i^{Exp} + v_{Exp,i} \quad (4.3b)$$

Finally, the probability of receiving foreign direct investment equation depends on firm level productivity (TFP), the investment climate and by the fact that certain firm's exports;

$$y_{it}^{FDI} = \rho_{FDI} + a_{FDI,i} + \rho_P \log TFP_{it} + \rho_{Exp} y_{it}^{Exp} + \rho'_{DR} D_r + \rho'_{Ds} D_j + \rho'_{DM} D_m + \rho'_{DT} D_t + \varepsilon_{FDI,it} \quad (4.4a)$$

$$a_{FDI,i} = \rho'_{IC} IC_i^{FDI} + \rho'_C C_i^{FDI} + v_{FDI,i} \quad (4.4b)$$

Notice that since the variable y^r_{it}, with r = Exp or FDI, is a *binary random variable* taking only 0 and 1 values, then $P(y_{it}^r = 1/x) = E(y_{it}^r / x)$ then: a) the conditional probability is equal to the conditional expectation which is usually assumed to follow a Probit or a Logit model, and b) the conditional variance (heteroskedasticity) is equal to the product of the conditional probabilities of the two events. In general, the linear probability models (LPM) approximate well the Probit and Logit nonlinear models when the variables are evaluated close to their sample means. Since we are interested in the mean IC contribution relative to the mean values of the dependent variables of (4.1a) to (4.4a), we will concentrate only on linear probability specifications, like (4.3a) and (4.4a). The main advantage of the LPM is in its simplicity since the parameters of the explanatory variables of (4.3a) and (4.4a) measure the change in probability when one of the explanatory variables changes, holding the rest of the explanatory variables constant. This is important for the economic interpretation of the coefficients obtained in the empirical section.

By substituting the usually unobserved fixed effects components by their corresponding equation we can simplify the system of equations including productivity to:

$$\log TFP_{it} = \alpha_P + \alpha'_{IC} IC_i^P + \alpha'_C C_i^P + \alpha_{Exp} y_{it}^{Exp} + \alpha_{FDI} y_{it}^{FDI} + \alpha'_{DR} D_r + \alpha'_{Ds} D_j + \alpha'_{DM} D_m + \alpha'_{DT} D_t + (v_{P,i} + \varepsilon_{P,it}) \quad (4.5)$$

$$logL_{it} = \gamma_L + \gamma_P \log TFP_{it} + \gamma_w \log W_{it} + \gamma_{Exp} y_{it}^{Exp} + \gamma_{FDI} y_{it}^{FDI} + \gamma'_L IC_i^L + \gamma'_C C_i^L + \gamma'_{DR} D_r + \gamma'_{Ds} D_j + \gamma'_{DM} D_m + \gamma'_{DT} D_t + (v_{L,i} + \varepsilon_{L,it}) \quad (4.6)$$

$$\log W_{it} = \beta_W + \beta_P \log TFP_{it} + \beta_{Exp} y_{it}^{Exp} + \beta_{FDI} y_{it}^{FDI} + \beta_{IC}' IC_i + \beta_C' C_i + \beta_{DR}' D_r + \beta_{Ds}' D_j + \beta_{DM}' D_m + \beta_{DT}' D_t + (v_{W,i} + \varepsilon_{W,it}) \quad (4.7)$$

$$y_{it}^{Exp} = \delta_{Exp} + \delta_P \log TFP_{it} + \delta_{FDI} y_{it}^{FDI} + \delta_{IC}' IC_i^{Exp} + \delta_C' C_i^{Exp} + \delta_{DR}' D_r + \delta_{Ds}' D_j + \delta_{DM}' D_m + \delta_{DT}' D_t + (v_{Exp,i} + \varepsilon_{Exp,it}) \quad (4.8)$$

$$y_{it}^{FDI} = \rho_{FDI} + \rho_P \log TFP_{it} + \rho_{Exp} y_{it}^{Exp} + \rho_{IC}' IC_i^{FDI} + \rho_C' C_i^{FDI} + \rho_{DR}' D_r + \rho_{Ds}' D_j + \rho_{DM}' D_m + \rho_{DT}' D_t + (v_{FDI,i} + \varepsilon_{FDI,it}) \quad (4.9)$$

The composite error terms of each equation of the system have three terms, says $\omega_{it} = \gamma_{r,i} + v_{r,i} + u_{r,it}$ with *r=P, L, W, Exp* and *FDI*. The firm fixed effects ($\gamma_{r,i}$) are approximated by the set of observed time-invariant, firm level IC and C variables. The remaining unobserved firm effects are assumed to be independently distributed of IC and C variables, therefore what remains are random effects ($v_{P,i}$). Therefore, we assume that the error terms ($v_{r,i}+\varepsilon_{r,j,it}$) are uncorrelated with all the explanatory variables of each equation *r*, where *r=P, Exp, FDI, W* and *L*. However, for certain explanatory variables this exogeneity condition is not satisfied. The endogeneity of certain IC variables induces a correlation between those IC variables and the errors ($v_{r,i}+\varepsilon_{r,j,it}$) of the system of equations (4.5) to (4.9) and creates simultaneous equation biases and inconsistencies in least squares estimators; like pooling OLS or in random effects (RE) estimators. This correlation is in general mitigated by replacing those plant-level IC variables by their region-industry averages ($\overline{IC}_j$). However, for some other explanatory variables like productivity, wages, exports and FDI, the endogeneity is intrinsic due to the simultaneous structure of the system of equations. Therefore, we estimate each equation by instrumental variables (IV) techniques (2SLS) using heteroskedasticity-robust standard errors. We could have used 3SLS, which is more efficient than 2SLS under correct specification. However, since with system of equations estimation techniques the misspecification of one equation affects the whole system, we believe that the results from 2SLS are more robust.

Provided that we are instrumenting the productivity (TFP) variable in the employment, real wages, exports and FDI equations using instruments from the investment climate survey, it is very convenient to specify a number of rules to choose the list of instruments, etc. First, estimation of the system of equations (4.5) to (4.9) by IV techniques is done equation by equation. Productivity equation is at the core of this process and it is estimated seeking robust procedures of Escribano and Guasch (2005 and 2008). Once we have obtained robust IC and C coefficients for different productivity (TFP) measures, we use the set of significant explanatory variables to instrument productivity in the rest of equations. Notice that some of these variables will be used as included instruments, while many other will be excluded instruments as they may appear as explanatory variables in other equations.

The next step is to obtain a preliminary specification for the remaining equations of the system by OLS with robust standard errors. As in the productivity case, in order to avoid omitted variables problems, the selection of the model goes from the general to the specific. We start selecting the preliminary model from a set of more than 90 IC and C variables, industry, state and size dummies, productivity and a constant term (also real wages in the case of demand for labor equation).

Once we have a preliminary valid model for each equation of the system we start instrumenting productivity. We then remove instruments from the list of excluded instruments provided we want a partial R-squared –or '*Shea*' partial R-squared—as high as possible with the restriction that our model is not over-identified. To test the over-identification restrictions we use Hansen test, a robust to general heteroskedasticity variation of classical Sargan test. In addition we take into account the significance in the first stage estimates when removing instruments. We also remove instruments from the matrix of included instruments if in the process of IV selection some of them become insignificant.

A similar process is applied when we have to instrument any other simultaneous variable like real wages in demand for labor equation, or exports or FDI when they appear as significant explanatory variable in other equations. A good strategy that works well is to estimate first by OLS and then change to IV if we have the set of instruments, which in this case is given by the explanatory variables of the corresponding equation, excluding obviously those endogenous covariates. Then we proceed as in the productivity case, removing instruments, either included or excluded, according to the criteria mentioned before.

4.1 Identification of the system of equations

To discuss the identification issues underlying the system of equations proposed it is useful to apply matrix notation. The structural form of the system (4.5) - (4.9) is given by

$$\mathbf{A}\mathbf{y}_\mathbf{t} + \mathbf{B}\mathbf{x}_\mathbf{t} = \mathbf{u}_\mathbf{t} \qquad (4.10)$$

where $\mathbf{y}_\mathbf{t}$ is the 5×1 vector of observations of *dependent* variables (log-productivity, y_{it}^{Exp} and y_{it}^{FDI}, log-employment and log-wages); $\mathbf{x}_\mathbf{t}$ is the 140x1 vector of explanatory variables (IC_i, C_i, D_r, D_j and D_t); $\mathbf{u}_\mathbf{t}$ is the 5×1 vector of errors; $\mathbf{A}$ is a 5×5 matrix of coefficients of simultaneous *dependent* variables; $\mathbf{B}$ is a 5x164 matrix of coefficients corresponding to the exogenous/endogenous IC and variables.

In the system (4.5) - (4.9), we are imposing certain structure; for example that employment has no direct effect in any other equation of the system and that real wages only affects employment demand, after controlling for all IC and C variables. Therefore, we can explicitly write the first LHS term of (4.10) as;

$$\mathbf{Ay_t} \equiv \begin{pmatrix} 1 & -a_{P,Exp} & -a_{P,FDI} & 0 & 0 \\ -a_{Exp,P} & 1 & -a_{Exp,FDI} & 0 & 0 \\ -a_{FDI,P} & -a_{FDI,Exp} & 1 & 0 & 0 \\ -a_{L,P} & -a_{L,Exp} & -a_{L,FDI} & 1 & -a_{L,W} \\ -a_{W,P} & -a_{W,Exp} & -a_{W,FDI} & 0 & 1 \end{pmatrix} \begin{pmatrix} \log TFP_{it} \\ y_{it}^{Exp} \\ y_{it}^{FDI} \\ \log L_{it} \\ \log W_{it} \end{pmatrix} \equiv \begin{pmatrix} \log TFP_{it} - a_{P,Exp} y_{it}^{Exp} - a_{P,FDI} y_{it}^{FDI} \\ y_{it}^{Exp} - a_{Exp,P} \log TFP_{it} - a_{Exp,FDI} y_{it}^{FDI} \\ y_{it}^{FDI} - a_{FDI,P} \log TFP_{it} - a_{FDI,Exp} y_{it}^{Exp} \\ \log L_{it} - a_{L,P} \log TFP_{it} - a_{L,W} \log W_{it} - a_{L,Exp} y_{it}^{Exp} - a_{L,FDI} y_{it}^{FDI} \\ \log W_{it} - a_{W,P} \log TFP_{it} - a_{W,Exp} y_{it}^{Exp} - a_{W,FDI} y_{it}^{FDI} \end{pmatrix}.$$

The rank condition is a *necessary and sufficient condition* for the system (4.10) to be identified. To discuss whether the rank condition is satisfied, say, in the first equation, let $\boldsymbol{\alpha}'$ be the first row of $\mathbf{A}$ and $\boldsymbol{\beta}'$ the first row of $\mathbf{B}$. We may now partition these vectors into two components corresponding to the included ($\boldsymbol{\alpha}_1'$ and $\boldsymbol{\beta}_1'$) variables and the excluded ($\boldsymbol{\alpha}_2'$ and $\boldsymbol{\beta}_2'$) variables in the productivity equation such that $\mathbf{A} = \begin{bmatrix} \boldsymbol{\alpha}_1' & \mathbf{0} \\ \mathbf{A}_1 & \mathbf{A}_2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} \boldsymbol{\beta}_1' & \mathbf{0} \\ \mathbf{B}_1 & \mathbf{B}_2 \end{bmatrix}$, which allow us to construct the next matrix $\mathbf{D} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{B}_2 \end{bmatrix}$. By the rank condition, productivity equation is identified if $rank(\mathbf{D}) = 5-1$. The same holds for the rest of equations of the system. Thus, even if we have several exclusion restrictions in matrix $\mathbf{A}$ (in the productivity, wages and employment equations), nevertheless these restrictions are not enough to ensure that the rank condition is satisfied. For that, we force the coefficient of certain IC variables to be 0 prior to start estimating the system, for more details on extra identification issues see Escribano et al (2008b).

The empirical IC results based on 2SLS are included in Tables E.I to E.IV of the Appendix B. In all the cases we found evidence that TFP has a significant and positive impact on; employment demand, on real wages, and on the probabilities of exporting or receiving FDI. Notice that TFP is always significant even after controlling for IC and other C variables.

5. IC assessment on aggregate productivity (Olley and Pakes decomposition) and other measures of economic performance.

In the second part of the analysis, taking advantage of the robustness of the IC, and C elasticities estimated, we want to concentrate on the TFP measure that comes from the restricted Solow's residuals. Our aim is to evaluate the IC effects on average productivity and on allocative efficiency components of the Olley and Pakes (1996) decomposition (O&P) of aggregate productivity in levels (TFP) and on the mixed O&P decomposition (logTFP).

5.1 O&P decompositions: in levels and mixed.

The O&P decomposition of aggregate productivity in levels is,

$$TFP = \overline{TFP} + N\,\hat{\text{cov}}(s_{it}^{Y}, TFP_{it}) \,. \tag{5.1a}$$

Where TFP is aggregate productivity (TFP) (or weighted average productivity, where the weights are given by the share of sales), $\overline{TFP}$ is the sample average productivity and the last term is N times the sample covariance of the share of sales and firm level productivity; this last term is the allocative efficiency term describing the ability of the markets to reallocate resources from less to more productive establishments. Furthermore, we want to exploit the log-linear properties of the following mixed[23] O&P decomposition in order to obtain closed form O&P decompositions in terms of IC and C variables,

$$\log TFP = \overline{\log TFP} + N\,\hat{\text{cov}}(s_{it}^{Y}, \log TFP_{it}) \,. \tag{5.1b}$$

Expressions (5.1a) and (5.1b) can be easily applied by industry, state, size, age or for the whole sample. The results of the decomposition by states and at country level in levels and mixed decomposition are in Figures 2 and 3 of appendix B.

5.2 IC effects on productivity measure in the terms of the mixed O&P decomposition.

The useful additive property of equation (3.3) in logarithms, allow us to obtain an exact closed form solution of the decomposition of aggregate log productivity according to equation (5.1b). Following Escribano et al. (2008a), we can express aggregate log productivity as a weighted sum of the average values of the IC, C, dummy D variables, the intercept and the productivity average residuals ($\bar{\hat{u}}$) from (3.3); and, the sum of the *covariances* between the share of sales and investment climate variables IC, C, dummies D and the productivity residuals ($\hat{u}$).

$$\begin{aligned}\log TFP = {} & \hat{\alpha}'_{IC}\overline{IC}_{P} + \hat{\alpha}'_{C}\bar{C}_{P} + \hat{\alpha}'_{DR}\bar{D}_{r} + \hat{\alpha}'_{Ds}\bar{D}_{j} + \hat{\alpha}'_{DM}\bar{D}_{m} + \hat{\alpha}'_{DT}\bar{D}_{t} + \hat{\alpha}_{p} + \bar{\hat{u}}_{it} \\ & + N\hat{\alpha}'_{IC}\hat{\text{cov}}(s_{it}^{Y}, IC_{P,i}) + N\hat{\alpha}'_{C}\hat{\text{cov}}(s_{it}^{Y}, C_{P,i}) + N_{q}\hat{\alpha}'_{Ds}\hat{\text{cov}}(s_{it}^{Y}, D_{j}) + N\hat{\alpha}'_{DR}\hat{\text{cov}}(s_{it}^{Y}, D_{r}) \\ & + N\hat{\alpha}'_{DT}\hat{\text{cov}}(s_{it}^{Y}, D_{t}) + N\hat{\alpha}'_{DM}\hat{\text{cov}}(s_{it}^{Y}, \bar{D}_{m}) + N\hat{\alpha}'_{DT}\hat{\text{cov}}(s_{it}^{Y}, \bar{D}_{t}) + N\hat{\text{cov}}(s_{it}^{Y}, \hat{u}_{it})\end{aligned} \tag{5.2}$$

[23] It is called mixed Olley and Pakes (O&P) decomposition because in the original O&P decomposition both TFP and the share of sales were in levels while now TFP in (5.1b) is in logs, (log P).

where the set of estimated parameters used comes from the two-step TFP estimation, having the restricted Solow's residual as dependent variable of the regression equation (3.3).

The contributions of IC variables to aggregate log-TFP of equation (5.2) can be computed for the whole sample or by industry/sector, state, size, etc. In particular, we compute the IC contributions relative to aggregate productivity as follows;

$$\begin{aligned} 100 = \frac{100}{\log TFP}[& \hat{\alpha}'_{IC}\overline{IC}_P + \hat{\alpha}'_C \overline{C}_P + \hat{\alpha}'_{DR}\overline{D}_r + \hat{\alpha}'_{Ds}\overline{D}_j + \hat{\alpha}'_{DM}\overline{D}_m + \hat{\alpha}'_{DT}\overline{D}_t + \hat{\alpha}_p + \overline{\hat{u}}_{it} \\ & + N\hat{\alpha}'_{IC}\hat{\text{c}}\text{ov}(s^Y_{it}, IC_{P,i}) + N\hat{\alpha}'_C \hat{\text{c}}\text{ov}(s^Y_{it}, C_{P,i}) + N_q\hat{\alpha}'_{Ds}\hat{\text{c}}\text{ov}(s^Y_{it}, D_j) + N\hat{\alpha}'_{DR}\hat{\text{c}}\text{ov}(s^Y_{it}, D_r) \\ & + N\hat{\alpha}'_{DT}\hat{\text{c}}\text{ov}(s^Y_{it}, D_t) + N\hat{\alpha}'_{DM}\hat{\text{c}}\text{ov}(s^Y_{it}, \overline{D}_m) + N\hat{\alpha}'_{DT}\hat{\text{c}}\text{ov}(s^Y_{it}, \overline{D}_t) + N\hat{\text{c}}\text{ov}(s^Y_{it}, \hat{u}_{it})]. \end{aligned} \quad (5.3)$$

There are several advantages of using equation (5.3). First, we can compare net contributions by isolating the impact of IC variables from the impact of industry dummies, the intercept, and the residuals. Second, we can split the total effect on aggregate productivity in the part explained only by IC, and C variables (demeaned *logTFP*), and the proportion is due to the rest; constant term, industry dummies and so on. The empirical results of decomposition (5.3) are in Table C.II.

We could also get rid of the different directional effects (positive or negative) of the various IC effects by simply computing the percentage contributions in absolute value. This slightly modification allow us to do direct comparisons of the IC absolute percentage contributions (or weight of each IC variable relative to the total weight of other IC variables) to aggregate log-productivity, to average log-productivity and to the allocative efficiency term. The results are in Figure 4.1.

5.3 Simulations based on the IC effects on the O&P decomposition of TFP.

So far, we have exploited the linear properties of the logarithm form of the mixed O&P decomposition of TFP. However, the original O&P decomposition was done in terms of TFP and the share of sales (in levels). Therefore the O&P decompositions is capturing also nonlinear relations between market shares and IC variables coming from (5.1a) and equation (3.3). To know to what extent these nonlinear terms are affecting this relation, we perform simulation experiments[24] on INF, IC, and C variables, and evaluate the consistency of the results with the ones obtained from the previous mixed O&P decomposition- see (5.3). The IC simulations are done variable by variable (one at a time) keeping the rest of the variables constant; that is, we propose a scenario in which one of the IC variables experiment a 20 percent improvement in all the establishments. We compute the corresponding rate of change of aggregate productivity, average productivity and allocative efficiency caused by such improvement. We repeat the same experiment for the rest IC and C

[24] We are indebted to Ariel Pakes for this suggestion.

variables, and, for comparative purposes, we also evaluate the relative IC effect by group of IC variables.

The resulting simulations of a 20% improvement in IC variables are in Figure 4.2 of appendix B. A comparison between the simulations and the IC absolute percentage contributions are in Figure 4.3 in the same appendix.

5.4 IC evaluation on the sample means of employment and wages, on the probability of exporting and on the probability of receiving FDI.

The objective now is to measure the *partial direct effect* of each IC variable on each dependent measuring economic performance from the system of equations (4.5)-(4.9), at different aggregation levels (aggregate level, by sector, by region, by size of the firm, by age of the firm, etc.). For that purpose, we evaluate the impact of the average IC variable on the sample average values of the dependent variables of the system. In what follows, we substitute all the unknown parameters of the system (4.5) to (4.9) by their corresponding 2SLS estimated values.

The labor demand and the wage equations evaluated at the sample means and in relative terms are,

$$100 = \frac{100}{\overline{\log L_t}}[\hat{\gamma}_L + \hat{\gamma}_P \overline{\log TFP}_t + \hat{\gamma}_w \overline{\log W}_t + \hat{\gamma}_{Exp} \overline{y}_t^{Exp} + \hat{\gamma}_{FDI} y_t^{FDI} + \hat{\gamma}'_L \overline{IC}^L + \hat{\gamma}'_C \overline{C}^L + \hat{\gamma}'_{DR} \overline{D}_r + \hat{\gamma}'_{Ds} \overline{D}_j + \hat{\gamma}'_{DM} \overline{D}_m] \qquad (5.4)$$

$$100 = \frac{100}{\overline{\log W}_t} \hat{\beta}_W + \hat{\beta}_P \overline{\log TFP}_t + \hat{\beta}_{Exp} y_t^{Exp} + \hat{\beta}_{FDI} y_t^{FDI} + \hat{\beta}'_{IC} \overline{IC}^W + \hat{\beta}'_C \overline{C}^W + \hat{\beta}'_{Ds} \overline{D}_j + \hat{\beta}'_{Ds} \overline{D}_j + \hat{\beta}'_{DR} \overline{D}_r + \hat{\beta}'_{DM} \overline{D}_m \, . \qquad (5.5)$$

Since y_{it}^{Exp} and y_{it}^{FDI} are binary variables, evaluating the impact at the sample mean implies the evaluation on the *probability (frequency) of exporting and receiving FDI*, respectively. In particular equations (3.8) and (3.9) relative to the frequency of exporting and receiving FDI becomes

$$100 = \frac{100}{\hat{P}(Exp_t \succ 0)} \hat{\delta}_{Exp} + \hat{\delta}_P \overline{\log TFP}_t + \hat{\delta}_{FDI} y_{it}^{FDI} + \hat{\delta}'_{IC} \overline{IC}^{Esp} + \hat{\delta}'_C \overline{C}^{Exp} + \hat{\delta}'_{Ds} \overline{D}_j + \hat{\delta}'_{DR} \overline{D}_r + \hat{\delta}'_{DM} \overline{D}_m \qquad (5.6)$$

$$100 = \frac{100}{\hat{P}(FDI_t \succ 0)} \left[\hat{\rho}_{FDI} + \hat{\rho}_P \overline{\log TFP}_t + \rho_{Exp} \overline{y}_t^{Exp} + \hat{\rho}'_{IC} \overline{IC}^{FDI} + \hat{\rho}'_C \overline{C}^{FDI} + \hat{\rho}'_{Ds} \overline{D}_j + \hat{\rho}'_{DR} \overline{D}_r + \hat{\rho}'_{DM} \overline{D}_m \right] \qquad (5.7)$$

The results of equations (5.4) to (5.7) are in Figures 5.1 to 5.10 of appendix B.

6. Empirical results.

6.1 Key Results on Productivity

Figure 5.5 in its first column compare the relative importance of groups of IC variables in terms of contributions to average log-productivity for the case of total ICT sector analysis and Figure 5.10 shows the same contribution in the Software and ITES sub-sector analysis.

Infrastructures group is the most important group in both analyses. For total ICT sector, infrastructure factors represent 72.2% of the whole contribution of IC and C variables to average log-productivity; within these factors, productivity is affected (see Figure 4.1) by having an own electric generator, having a high speed internet connection, using internet to deliver services to this establishment's clients and by the number of days waiting for an electric connection. All contributions are relevant, although the number of days waiting for an electric connection is the variable with the highest contribution. In the case of the analysis of Software and ITES sub-sector, Infrastructures group relative importance with respect to all IC variables in productivity equation is 58.5%, as Figure 5.10 shows. Figure 4.6 breaks down this percentage in key factors: having an own electric generator, having a high speed internet connection, suffering unavailability of quality internet connection and using internet to deliver services to this establishment's clients. Dummy for high speed internet connection has a high contribution too, but the largest contributions are given by the dummy for use internet to deliver services to this establishment's clients.

In the two analyses, *other control variables* is the second group in order of importance. As figures 5.5 and 5.10 show, its contributions are of 13.4% and 24.9%, respectively. The only variable with significant impact within this group is dummy for incorporate company, being one of the highest individual contributions of all the IC variables, especially in the case of Software and ITES sub-sector. (See figures 4.1 and 4.6)

Quality, innovation and labor skills factors represent 6.8% of the whole contribution of IC and C variables to average log-productivity in the case of whole ICT sector. The factors of this group affecting productivity are: r&d expenditures and dummy for external training with similar and small contributions in ICT sector as Figure 4.1 shows. For the analysis of Software and ITES sub-sector, the contribution of this group is the lowest, with a relative impact of only 2.3%, as Figure 5.10 illustrates. The two significant variables are dummy for receiving royalty payments and the percentage of royalty payments received. (See Figure 4.6)

Red tape, informality and others is the third group in order of importance representing 7.5% of the whole contribution of IC variables to average log-productivity (see Figure 5.19) for the case of Software and ITES sub-sector. In the ICT sector analysis is the fourth group in order of importance representing 5.8% of the total contributions (see Figure 5.5). Within this group three variables are significant in each analysis. Dummy for conflicts in courts, manager's time spent in bureaucratic issues and number of inspections for total ICT sector and dummy for conflicts in courts, manager's time spent in bureaucratic issues and sales never repaid for Software and ITES sub-sector. As Figures 3.3 and 6.3 show, the largest contributions to average log-productivity within this group come from manager's time spent in bureaucratic issues in total ICT and sales never repaid in Software and ITES.

Finally, *finance and corporate governance* group has a relative impact on average log-productivity of only 2.1% (ICT sector) and of 6.7% (Software and ITES sub-sector) as Figures 5.5 and 5.10 show. The only two variables significant in both analyses are working capital financed by state-owned banks and dummy for overdraft, being the contribution of having an overdraft higher in the two cases. (See Figure 4.1 and Figure 4.6)

Olley and Pakes Decomposition

We focus now on the decomposition of the allocative efficiency term (or covariance term) of the mixed Olley and Pakes decomposition for the case of total ICT sector analysis. Column *contributions* of the *Allocative Efficiency* section of Figure 4.3 shows the relative impact of each group of IC and C variables on this term at the aggregate level. The main group affecting the allocative efficiency is *infrastructures* representing 56.4% of the whole contribution of IC and C variables to the allocative efficiency. Next group is *red tape, informality and others* being its weight 12.9%, followed by *other control variables* which weight is 12.6%. The relative contributions of *quality, innovation and labor skills* and *finance and corporate governance* are 10.4% and 7.7% respectively. Figure 4.1 shows what variables have the largest contributions on the allocative efficiency, these variables are: wait for an electric supply, dummy for own generator, dummy for incorporate company and manager's time spent in bureaucratic issues. These contributions have a straightforward interpretation, the larger and positive (see Table C.2) the contribution of an IC variable the more productions is causing that variable going from less efficient firms to more efficient ones, vice versa if the effect is negative.

For Software and ITES sub-sector (Figure 4.8) the main group affecting the allocative efficiency is *infrastructures* representing 33.4% of the whole contribution of IC and C variables to the allocative efficiency. Next group is *finance and corporate governance* being its weight 26.9%, followed by *red tape, informality and others* which weight is 18.7%. The relative contributions of *other control* *variables* and *quality, innovation and labor skills* are 12.4% and 8.5% respectively. Figure 4.6 shows what variables have the largest contributions on the allocative efficiency, these variables are: dummy for overdraft, dummy for own generator and dummy for incorporate company.

Exporters versus non-exporters

Figures 4.4 and 4.5 show the individual contribution of each IC variable to mixed Olley and Pakes decomposition for the cases of ICT sector and Sofware and ITES sub-sector, respectively.

For the exporter firms the highest contributions on average productivity came from infrastructure factors: dummy for high speed internet connection, dummy for internet uses to deliver services and wait for an electric supply. However the largest impacts on the other component of Olley and Pakes decomposition, allocative efficiency, come from: R&D expenditures, dummy for conflicts in courts and number of inspections. (See Figure 4.4)

If the firm does not export (Figure 4.5), the main variables affecting the average productivity are: dummy for internet uses to deliver services, wait for an electric supply and number of inspections. Having dummy for high speed internet connection and dummy for incorporate company large contributions. The most relevant impacts of IC variables on allocative efficiency component come from: wait for an electric supply, number of inspections, R&D expenditures and dummy for incorporate company.

6.2 Key Results on Employment

Figure 5.5 (for ICT sector) and Figure 5.10 (for Software and ITES sub-sector) in its second column compares the relative importance of groups of IC and C variables in terms of contributions to average log-employment at the aggregate level.

In both analyses the contribution of *productivity* to average log-employment is similar, 8.8% (ICT sector) and 11.7% (Software and ITES sub-sector) as Figure 5.5 and Figure 5.10 show.

Similar conclusions we can obtain in the case on *real wages* impact on average log-employment. For total ICT sector *real wages* represents 38.1% of total IC and C contributions being the first group in order of importance. *Real wages* has the largest relative impact of the whole contribution of IC and C variables to average log-employment representing the 31% for the software and ITES analysis.

Regarding *infrastructures* its relative weight in Figure 5.5 is only 5.3% and in Figure 5.10 is only 4.8%. The infrastructure factors affecting the demand of employment are (see Figures 5.1 and 5.6) having own generator and number of unavailability of quality internet connection (in ICT sector) and having own generator (in Software and ITES sub-sector).

In what refers to *red tape, informality and others* group its relative weight in Figure 5.5 is 15.9% and in Figure 5.10 is 12%. Figure 5.1 describes the factors of this group affecting employment in ICT sector analysis: dummy for conflicts in courts, number of weeks to resolve overdue payments and security cost. The largest contribution of this group is given by security cost variable. For Software and ITES sub-sector (see figure 5.6) the factors are: dummy for conflicts in courts, number of weeks to resolve overdue payments and manager's time spent in bureaucratic issues. The largest contribution comes from number of weeks to resolve overdue payments.

Finance and corporate governance weight is 6.7% (ICT sector) and 11% (Software and ITES sub-sector). Three variables have impact within this group in the two analyses: the percentage of working capital financed by private commercial banks, the percentage of rent land and dummy for acquiring additional land for ICT sector; and the percentage of working capital financed by private commercial banks, having a loan and the percentage of rent land for Software and ITES sub-sector. Figure 5.1 and Figure 5.6 show that the percentage of rent land has the largest contribution on employment demand in both cases.

The relative weight of *quality, innovation and labor skills* group is 1.9% in total ICT sector and 5.2% in Software and ITES sub-sector. The factors of this group involving the demand of employment are presented in Figures 5.1 and 5.6 and they are commented in what follows: dummy for quality certification and dummy for receiving royalty payments are significant in the total ICT sector analysis; and dummy for quality certification and dummy for external training in the Software and ITES sub-sector analysis. The largest contributions of this group of variables come from dummy for quality certification (total ICT) and from dummy for external training (Software and ITES).

The second largest relative impact comes from *other control variables*, representing 23.3% of the whole contribution of IC and C variables to average log-employment in the case of total ICT sector. For Software and ITES sub-sector, this group is the second group in order of importance with a relative contribution of 24.4%. The factors of this group involving the demand of employment are: the age of the firm, percentage of direct exports and dummy for exporting processing zone in the case of ICT sector (see Figure 5.1), and only the age of the firm and the percentage of direct exports for Software and ITES sub-sector (see Figure 5.6). Always the largest contribution comes from the age of the firm.

6.3 Key Results on Real Wages

Third column of Figure 5.5 and of Figure 5.10 illustrates the relative importance explaining average log-wages of each group of IC and C variables. Real wages are closely and positively related with productivity, real wages increase as firms become more productivity. Productivity has an important contribution with a percentage of 29.1% in total ICT sector and with a percentage of 23.3% in the other analysis. Productivity relative importance becomes even larger when compared individually with other IC factors as Figures 5.2 and 5.7 show.

The relative weight of *infrastructures* group is only 3.4% in total ICT sector (see Figure 5.5); however in Software and ITES sub-sector the impact represent 45.5% of the total contributions being the first group in order of importance. (See figure 5.10). Only one variable is significant within this group, as Figure 5.2 and Figure 5.7 show; dummy for unavailability of quality internet connection (ICT sector) and number of power outages (Software and ITES sub-sector).

The relative weight of *red tape, informality and others* group in Figure 5.5 is 11.6% and in Figure 5.10 is 12.8%. Figure 5.2 highlights which concrete factor of this group have effect on wages in the total ICT sector analysis, the security cost with an important contribution. For Software and ITES sub-sector analysis, the only variable with effect on real wages is sales never repaid.

Finance and corporate governance group has a relative weight in total ICT sector of 5.4% and in Software and ITES sub-sector of 1.4%. Figure 5.2 and Figure 5.10 list the factors of this group; the percentage of working capital financed by family and friends and dummy for acquired or attempted to acquire additional land (ICT sector) and the the percentage of working capital financed by family and friends (software and ITES sub-sector).

Quality, innovation and labor skills group represents only 2.3% of the whole contribution of IC and C variables to average log-wage in Figure 5.10, but 43.4% in Figure 5.5. Specifically, the only factor of this group that has an effect on wages is dummy for R&D in Software and ITES sub-sector, being its percentage contribution very small. However in the case of total ICT sector two variables have a significant contribution on real wages: dummy for R&D and the percentage of staff for which English is critical, with a very important contribution (see Figure 5.2), the largest of the individual impacts.

Finally, *other control variables* group represents 7% of total IC and C variables contribution on average log-wages for total ICT sector. Figure 5.2 shows the specific factor of this group: the share of direct exports. For Software and ITES sub-sector analysis, *other control variables* group has a relative weight of 14.8% (see Figure 5.10). Only one variable is significant within this group in this analysis, the age of the firms. (See Figure 5.7)

6.4 Key Results on the Probability of Exporting

As Figure 5.5 in its fourth column shows; productivity explains 17.2% of the whole impact of IC and C variables on the probability of exporting for total ICT sector. From Figure 5.10 it is clear that the productivity impact on the probability of exporting in Software and ITES sub-sector is larger, being its relative weight of 31.2% and when compared individually with other IC factors productivity become even more important, being only exceeded by one IC variables in terms of percentage contributions to the probability of exporting. (See Figure 5.8)

Regarding *infrastructures*, its relative weight in Figure 5.5 an in Figure 5.10 is 15.9%. Improved infrastructures make easier to export. Having own generator or not having enough hotel accommodation have impact on probability of exporting in total ICT sector analysis. For Software and ITES sub-sector analysis the significant variables are: dummy for internet uses to connection to make purchases and dummy for hotel accommodation constrain. From Figures 5.3 and 5.8 we are able to identify which are the factors with the largest impact on the probability of exports. In both cases the largest contribution comes from dummy for hotel accommodation constrain variable.

The first group in order of importance in both analyses is *red tape, informality and others*, its relative weight in Figure 5.5 is 30.6% and in Figure 5.10 is 37%. Within this group the factors with impact on the probability of exporting in total ICT sector are: dummy for conflicts in courts and the percentage of sales never repaid. And in Software and ITES sub-sector are: dummy for conflicts in courts and security cost. The largest impact is given by sales never repaid variable (ICT sector), as Figure 5.3 shows; and by security cost variable (software and ITES sub-sector), as Figure 5.8 shows.

Finance and corporate governance factors have a relative weight in Figure 5.5 of only 1.9% and in Figure 5.10 of only 1.2%. Two variables within this group affect on export equation in each case: the percentage of working capital financed by informal sources and the percentage of new investment financed by new equity (ICT sector); and the percentage of working capital financed by family and friends and the percentage of working capital financed by informal sources (software and ITES sub-sector). All percentage contributions are very small.

In what refers to *quality, innovation and labor skills* group its relative weight in Figure 5.5 is 11.8% and in Figure 5.10 is 3%. Figure 5.3 describes the factors of this group affecting the probability of exporting in ICT sector analysis: the percentage of staff for which English is critical and dummy for external training. The largest contribution of this group is given by staff for which English is critical variable. For Software and ITES sub-sector (see figure 5.8) the factors are: dummy for quality certification and royalties received.

Other control variables group weight in Figure 5.5 is 22.6% and in Figure 5.10 is 11.7%. Within this group, the age of the firm and dummy for Export Processing Zone have impact on probability of exporting, coming the largest contribution from age variable (ICT sector). In Software and ITES sub-sector analysis, only the age of the firm has impact on probability of exporting.

6.5 Key Results on the Probability of Receiving Foreign Direct Investment

We now focus on the results of the foreign direct investment equation. Last column of Figure 5.5 shows that productivity is a key factor affecting FDI decisions; its weight in this figure is 43.6%. Its effect is positive, meaning that more productivity implies more probability of receive FDI. From Figure 5.4 it is clear that productivity has the largest contribution to the probability of receiving FDI among all IC and C variables in the case of total ICT sector. However the relative weight of productivity contribution on the probability of receiving FDI in Software and ITES is smaller, being 32.8% and the second group in order of importance. (See Figure 5.10)

Infrastructures group has a different behaviour in each case. In the total ICT sector analysis, its relative weight is very small (2.6%) and only dummy for own transport for workers variable is significant (see Figure 5.4). However, as Figure 5.10 shows, *infrastructures* group is the first group in order of importance with a relative weight of 33.8%. The infrastructures factors that affect the probability of receiving FDI are the number of power outages and dummy for own transport for workers. The largest contribution comes from the number of power outages variable. (See Figure 5.9)

Red tape, informality and others factors have a relative weight in Figure 5.5 of only 1.8% and in Figure 5.10 of only 1.3%. One variable within this group affect on FDI equation in both cases: dummy for cyber crime (see Figure 5.4 and Figure 5.9

Finance and corporate governance group has a relative weight in Figure 5.5 of 12.8%. The finance factors that affect the probability of receiving FDI are: the percentage of firm's working capital financed with informal sources and rent land variable. The largest contribution of this group is given by rent land (see Figure 5.4). For the case of Software and ITES sub-sector, the relative contribution of this group is 9.9%, as Figure 5.10 shows. The same two variables have a relevant impact on the probability of receiving FDI. (See Figure 5.9)

Regarding *quality, innovation and labor skills* its relative weight in Figure 5.5 is 25.1% and in Figure 5.10 is 15.5%. Specifically, the factor of this group that has an effect on probability of receiving FDI is shown in Figure 5.4 and in Figure 5.9, this factor is the percentage of workforce for which English is critical for two analyses.

Finally, *other control variables* group has a weight in Figure 5.5 of 11.4% and in Figure 5.10 of 6.6%. The percentage of direct exports is the only variable with a significant impact on the probability of receiving FDI within this group. (See Figure 5.4 and Figure 5.9)

7. Conclusions.

Applying the methodology for ICs developed by Escribano and Guasch (2005, 2008) and Escribano et al (2008a and 2008b), we identified the main IC constrains affecting productivity and other economic performance measures of ICT firms in India. ICT has been the fastest growing sector and is the largest contributor to exports. Within this sector, India is especially successful in software sub-sector, being one of the world superpower. However, the level of productivity in the ICT sector is lower than its potential.

The strategy for the identification of IC effects is based in the robustness of the empirical regularities. We found IC elasticities and semi-elasticities with respect to productivity that are robust under alternative economic environments (assumptions). Given this robustness, for the evaluation of IC effect on economic performance we concentrate in only one set of IC parameters. The idea is to obtain empirical regularities that are reasonably robust (in terms of signs and magnitude) under alternative econometric conditions, even when these assumptions do not hold. Obviously, the results are not numerically identical among different specifications but the observed variation of the estimates and significance is reasonable and gives more credibility to the empirical results obtained.

The identification and posterior assessment is not a straightforward task due to the numerous methodological difficulties we have encountered. To list a few endogeneity of regressors, productivity (TFP) measures, selection of the relevant model, simultaneous effects and low quality of the database (missing observations, outliers, etc.) have been addressed with the ICSs of India. We believe that the empirical regularities observed allow us to obtain a valuable insight on which are the main areas of reform regarding the investment climate.

Figure 4.3 summarizes the results obtained for productivity. Both IC percentage contributions and simulations reveal the important role of the infrastructures on productivity, particularly from the low quality of the supply of power and the internet connection. The total absolute IC contribution of the infrastructure group to aggregate productivity is around 70%.

Aggregate productivity is also associated with a number of red tape, informality and others variables, mainly related with bureaucracy and courts. Although the final absolute contribution of the red tape, informality and others variables is only 3.3%. An important contribution comes from the other control variables group. Its contribution (13.6) comes from the incorporated company variable, the only within this group with significant impact on productivity.

The association of finance and corporate governance is represented by the percentage of working capital financed by state-owned banks and the dummy for overdraft variable.

Finally, it is worth mentioning the contribution of quality, innovation and labor skills to aggregate productivity. R&D expenditures and having external training program contribute in overall with almost 14% to aggregate productivity.

Figure 4.8 shows the same study but dropping telecom services and Media and Entertainment industries. The most significant difference is the loss of importance of the infrastructures group and the gain on contribution of the variables within the group of finance and corporate governance.

The differences between exporter firms and non-exporter firms are illustrated in figures 4.4 and 4.5. Both groups of firms are very influenced by infrastructures factors. The main difference is the larger impact of R&D expenditures (overall on allocative efficiency component) for exporter firms. The contributions of number of inspections and dummy for incorporate company are higher in the case of non-exporter firms.

The IC contributions to the sample means of all the economic performance measures considered in this paper are summarized in Figure 5.5. The infrastructure group appears to be especially important for productivity and for the probability of exporting. The weight of the red tape, informality and others group is larger than 10% in employment and real wages and larger than 30% in exporting equation. The finance and corporate governance group only represents a significant contribution on the probability of receiving FDI equation. Quality, innovation and labor skills factors have more relative importance in real wages, exports and FDI. The Wage contribution to employment is 38.1%, and productivity (TFP) is a key factor in wages and FDI equations, contributing with 29% and 36% respectively.

The main differences with the results for Software and ITES sub-sector (Figure 5.10) are in the higher impact of the productivity on exports and in the large contribution of infrastructures group on real wage and FDI.

The investment climate is an important factor affecting the potential productivity growth of the firms in the sector. Limited skills in the workforce, electricity shortages and corruption are the biggest obstacles for exporters, whereas corruption, electricity shortages and access to land are the biggest obstacles for non-exporters. The perceptions of the managers are supported by econometric analysis.

References

Ackerberg D. A. and K. Caves (2003). "Structural Identification of Production Functions". *Working Paper,* University of California, Los Angeles.

Alexander, W. Robert J., Bell, John D., and Knowles, Stephen (2004). "Quantifying Compliance Costs of Small Businesses in New Zealand". Discussion paper, University of Otahgo.

Barro R. J. and X. Sala-i-Martin (2004). Economic Growth. The MIT Press, second edition, Cambridge, Massachusetts.

Bartelsman E. J. and M. Doms (2000). "Understanding Productivity: Lessons from Longitudinal Microdata". *Journal of Economic Literature*, Vol. 38, Nº 3, 569-594.

Bosworth, Barry and Susan Collins (2003). "The Empirics of Growth: An Update". The Brookings Institution. Washington, D.C. Processed.

Cecchini S. And C. Scott (2003). "Can Information and Communications Technology Applications Contribute to Poverty Reduction? Lessons from rural India". IOS Press.

Cole, H. L., L. E. Ohanian, A. Riascos and J. A. Schmitz Jr. (2004). "Latin America in the Rearview Mirror". National Bureau of Economic Research WP #11008, December.

Diewert W. Erwin and Alice O. Nakamura (2002) "The Measurement of Aggregate Total Factor Productivity Growth". J. J. Heckman and E. E. Leamer (eds.). Handbook of Econometrics, Vol. 6, forthcoming.

Djankov, S. and Hoekman, B., (2000). Foreign Investment and Productivity Growth in Czech Enterprises, *World Bank Economic Review*, 14, 49-64

Djankov, Simeon, La Porta, Rafael, Lopez-de-Silanis, Florencio, and Shleifer, Andrei (2002). "The Regulation of Entry". *Quarterly Journal of Economics* 117, February 2002. 1-37.

Escribano Alvaro and J. Luis Guasch (2005); "Assessing the Impact of the Investment Climate on Productivity using Firm Level Data: Methodology and the Cases of Guatemala, Honduras and Nicaragua," *World Bank Policy Research Working Paper* # 3621, The World Bank, June.

Escribano A. and J. L. Guasch (2008); "Robust Methodology for Investment Climate Assessment on Productivity: Application to Investment Climate Surveys from Central America," *Working Paper # 08-19 (11)*, *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena. (2008a); "Investment Climate and Firm's Performance: Econometric and Applications to Turkey's Investment Climate Survey," Working Paper # 08-20 (12), *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena (2008b); "Investment Climate Assessment Based on Demeaned Olley and Pakes Decompositions: Methodology and Applications to Turkey's Investment Climate Survey," Working Paper # 08-20 (13), *Universidad Carlos III de Madrid*.

Escribano A., N. Peltier, L. Garrido and H. Sing (2005b). "The Impact of Infrastructure on Competitiveness in Latin America: A Firm Level Analysis Based on Investment Climate Assessments". *Paper presented at the Joint Conference of the Interamerican Development Bank (IADB) and The World Bank (WB)*, June 6-7.

Escribano, A. and J. Pena (2008); "Empirical Econometric Evaluation of Alternative Methods of Dealing with Missing Values in Investment Climate Surveys," mimeo, *Universidad Carlos III de Madrid*.

Escribano, A., and R. Stucchi (2008). "Catching Up in Total Factor Productivity through the Business Cycle: Evidence from Spainish Manufacturing Surveys," Working Paper #08-51, Universidad Carlos III de Madrid.

Foster L., J. Haltiwanger and C.J. Krizan (1998). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence". *NBER Working Paper* W6803.

Guasch, J. Luis. (2004). "An Assesment of Logistic Costs and of their Impact on Competitiveness", World Bank, Washington, DC.

Hall R. E. (1990). "Invariance Properties of Solow´s Productivity Residual". In Peter Diamond (ed.). Growth, Productivity, Employment. Cambridge: MIT Press, 1-53.

Haltiwanger, John (2002). "Understanding Economic Growth. The Need for Micro Evidence". *New Zealand Economic Papers* 36 (1), 33-58.

He, Kathy S., Morck, Randall and Yeung, Bernard (2003). "Corporate Stability and Economic Growth". William Davidson Working Paper No. 553.

Hulten Ch. R. (2001). "Total Factor Productivity: A Short Biography". In Ch. R. Hulten, E.R. Dean and M. J. Harper (eds.) New Developments in Productivity Analysis, The University of Chicago Press, 1-47.

Iyengar P. (2006). "India`s ICT Industry: Increasing in Global Visibility and Relevance. Gartner.

Jorgenson D.W. (2001). "Information Technology and the U.S. Economy". *The American Economic Review*, Vol. 91, 1, 1-32.

Jorgenson D.W., F. Gollop and B. Fraumeni (1987). Productivity and U.S. Economic Growth. Cambridge: Harvard University Press.

Kasper, Wolfgang, (2002), "Losing Sight of the Lodestar of Economic Freedom: A Report Card on New Zeland's Economic Reform", NZ Business Roundtable.

Kerr, Roger (2002). "The Quagmire of Regulation" NZ Business Roundtable.

Levinsohn J. and A. Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables". *Review of Economic Studies*, 70, 317-341.

Loayza, N.V., A. M. Oviedo and Luis Serven (2004). "Regulation and Macroeconomic Performance", World Bank, Washington DC.

McMillan, John, (1998), " Managing Economic Change: Lessons from New Zeland", *The World Economy 21,* 827-43.

McMillan, John (2004), "A Flexible Economy? Entrepeneurship and Productivity in New Zeland", Working Paper, Graduate Scholl of Business, Stanford University, Stanford, CA.

Olley G. S.and A. Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry". *Econometrica*, Vol. 64, 6, 1263-1297.

Organization for Economic Cooperation and Development (2001). *Businesses' Views on Red Tape,* Paris, OECD.

Prescot, Edward C. (1998) "Needed: A Theory of Total Factor Productivity". International Economic Review, 39, 525-552.

Rodrik, Dani, and Arvind Subramanian (2004). "From 'Hindu Growth to Productivity Surge: The Mystery of the Indian Growth Transition". Harvard University, Cambridge, Mass. Processed.

Solow R. M. (1957). "Technical Change and the Aggregate Production Function". The *Review of Economics and Statistics*, 39 (3), 312-320.

Veeramani, C., and G. Goldar. (2004). "Investment Climate and Total Factor Productivity in Manufacturing: Analysis of Indian States." Working paper #127, Indian Council for Research on International Economic Relations, New Delhi, India.

World Bank (2005): "World Development Report 2005: A Better Investment Climate for Everyone," *The World Bank and Oxford University Press*, Washington, D.C.

World Bank (2009): "Doing Business Report 2009: Comparing Regulation in 181 Economies," *The World Bank Group*, Washington DC, USA.

World Bank 2007. "Unleashing India`s Innovation: Toward Sustainable and Inclusive Growth". Washington, D.C. World Bank

World Bank 2009. "India`s Investment Climate: Voices of Indian Business". Washington, D.C. World Bank

World Economic Forum, The (2008)."Global Competitiveness Report"

Wilkinson, Bryce (2001). "Constraining Government Regulation". NZ Business Roundtable.

Appendix A: Definitions

I Production Function Variables[25]

Sales: Used as the measure of output for the production function estimation. Sales are defined as total annual sales.

Employment: Total number of permanent and temporal workers.

Total hours worked per year: Total number of employees multiplied by the average hours worked per year.

Materials: Total cost of electricity, transport for goods and workers and communication services.

Capital stock: Net book value of ICT hardware and equipment (including transport) and land, buildings and leasehold improvements.

User cost of capital: The user cost of capital is defined in terms of the opportunity cost of using capital; it is defined as a 15% of the net book value of capital stock.

Labor cost: Total expenditures on personnel.

II Dependent Variables in Equation Regressions and Linear Probability Models

Demand for Labor: Total number of permanent and temporal workers.

Real Wage: Real wage is defined as the total expenditures on personnel divided by the total number of permanent and temporal workers.

Export: Dummy variable that takes value 1 if exports are greater than 10%.

Foreign Direct Investment: Dummy variable that takes value 1 if any part of the capital of the firm is foreign.

III General Information at Plant Level

Sector classification: a) telecom services; b) software; c) ITES; d) media and entertainment services.

Regional classification: a) Andhra Pradesh; b) Delhi; c) Haryana; d) Karnataka; e) Maharashtra; f) Tamil Nadu; g) Uttar Pradesh.

Size classification: a) micro firms (< 10 employees); b) small firms (>=10 & <50); c) medium firms (>=50 & <100); d) large firms (>=100).

[25] All series figure in US dollars, data obtained from The World Bank.

Table A.I: Definitions of Investment climate (IC) variables of Infrastructures group.

Name of the variable	Description of the variable	Observations (Response rate %)
Days to clear customs to import	Average number of days to clear customs when importing.	71 (31)
Longest days to clear customs to import	Longest number of days to clear customs when importing.	67 (29.3)
Dummy for power outages	Dummy variable taking value 1 if the firm has suffered power outages in last year.	115 (50.2)
Number of power outages	Total number of power outages suffered by the firm during last year.	82 (35.8)
Duration of power outages	Average duration of power outages suffered in hours, conditional on the pant reports having power outages.	79 (34.5)
Wait for electricity supply	Number of days waiting to obtain an electricity supply, conditional on submit an electrical connection.	56 (24.5)
Dummy for gifts electric supply.	Gifts expected or requested to obtain an electrical connection, conditional on submit an electrical connection.	75 (32.8)
Dummy for own generator	Dummy variable taking value 1 if the firm has its own power generator.	228 (99.6)
Electricity from a generator	Percentage of the electricity used by the plat provided by an own generator.	109 (47.6)
Wait for phone connection	Number of days waiting to obtain a phone connection, conditional on submit a phone connection.	86 (37.6)
Dummy for gifts for phone connection	Gifts expected or requested to obtain a phone supply, conditional on submit a phone connection	117 (51.1)
Dummy for high speed internet connection	Dummy variable taking value 1 if the firm uses a high speed internet connection .	226(98.7)
Dummy for internet to communicate	Dummy variable taking value 1 if the firm uses internet connection to communicate whit its clients or vendors.	214(93.4)
Dummy for internet to make purchases	Dummy variable taking value 1 if the firm uses internet connection to make purchases for this establishment.	211(92.1)
Dummy for internet to deliver services	Dummy variable taking value 1 if the firm uses internet connection to deliver services to this establishment's clients	216(94.3)
Dummy for internet to do research or develop ideas on new services	Dummy variable taking value 1 if the firm uses internet connection to do research or develop ideas on new services.	213(93)
Dummy for unavailability of quality internet connection	Dummy variable taking value 1 if the firm has suffered unavailability of quality internet connection in the last year.	217(94.8)
Number of unavailability of quality internet connection	Total number of unavailability of quality internet connection suffered by the firm during last year.	171(74.7)
Dummy for international flights	Dummy variable taking value 1 if in the last year were there enough direct international flights to and from its city	102(44.5)
Dummy for business	Dummy variable taking value 1 if in the last year were there enough business class seats for direct international flights to and from its city.	159(69.4)
Dummy for accommodation constrain	Dummy variable taking value 1 if the hotel accommodation has been a constrain in the last year.	165(72.1)
Dummy for own transport for workers	Dummy variable taking value 1 if the firm offers transport to its workers.	193(84.3)

Table A.II: Definitions of Investment climate (IC) variables of Red tape, informality and others group.

Name of the variable	Description of the variable	Observations (Response rate %)
Dummy for conflicts in courts	Dummy taking value 1 if the plant has conflicts with clients with a court involved (conditional on having conflicts with clients with a third part involved).	226(98.7)
Sales not paid in agreed time	Percentage of sales not paid in agreed time.	203(88.6)
Sales never repaid	Percentage of monthly total sales to private customers that were never repaid.	86(37.6)
Overdue payments	Number of days to resolve overdue payments.	89(38.9)
Weeks to judgment	Number of weeks that took the court to come to judgment in the last conflict with clients (conditional on having conflicts with clients with a third part involved).	74(32.3)
Dummy for cyber crime	Dummy taking value 1 if the plant has experienced losses due to cyber criminal attempts.	224(97.8)
Number of cyber crime	Total number of cyber criminal attempts suffered during last year.	69(30.1)
Security cost	Security expenses as a percentage of annual total sales.	85(37.1)
Manager's time spent in bureaucratic issues	In typical week percentage of manager's time spent dealing with bureaucratic issues.	217(94.8)
Dummy for payments to speed up bureaucracy	Gifts or informal payments to public officials to “get things done” with regard to customs, taxes, licenses, regulations, services etc.	113(49.3)
Wait for a construction permit	Days waiting to obtain a construction permit (conditional on submit a construction permit).	49(21.4)
Dummy for gifts to obtain a construction permit	Gifts expected or requested to obtain a construction permit, conditional on submit a construction permit.	63(27.5)
Wait for an import license	Total days to obtain an import license, conditional on submit an import license.	52(22.7)
Dummy for gifts for import license	Gifts expected or requested to obtain an import license, conditional on submit an import license.	65(28.4)
Wait for an operating license	Days waiting to obtain a main operating license (conditional on submit a operating license).	57(24.9)
Dummy for gifts for operating license	Gifts expected or requested to obtain a operating license, conditional on submit a operating license.	62(27.1)
Number of inspections	Total number of inspections of tax officials received by the plant in last year.	56(24.5)
Dummy payments for contract with the government	Dummy that takes value 1 if firms operating in the same sector of the surveyed plant have to offer informal payments to obtain a contract with the government.	163(71.2)
Payments to obtain a contract with the government	Payments to obtain a contract with the government as a percentage of contract value.	163(71.2)

Table A.III: Definitions of Investment climate (IC) variables of Finance and corporate governance group.

Name of the variable	Description of the variable	Observations (Response rate %)
Largest shareholder	Percentage of firm's capital owned by the largest shareholder.	187(81.7)
Dummy for trade association	Dummy that takes value 1 if the plant belongs to any association or trade chamber.	224(97.8)
Working capital financed by internal founds	Percentage of firm's working capital financed with internal funds.	224(97.8)
Working capital financed by private commercial banks	Percentage of firm's working capital financed with funds from commercial banks.	224(97.8)
Working capital financed by state-owned banks	Percentage of firm's working capital financed with funds from state banks.	224(97.8)
Working capital financed by trade credit	Percentage of firm's working capital financed with credits from suppliers or customers.	224(97.8)
Working capital financed by family/friends	Percentage of firm's working capital financed with family/friends funds.	224(97.8)
Working capital financed by informal sources	Percentage of firm's working capital financed with funds from informal sources.	224(97.8)
New investments financed by internal founds	Percentage of investments in new fixed assets financed with internal funds.	129(56.3)
New investments financed by new equity	Percentage of investments in new fixed assets financed with equity, sole of stock.	129(56.3)
New investments financed by new debt	Percentage of investments in new fixed assets financed with debt.	129(56.3)
New investments financed by private commercial banks	Percentage of investments in new fixed assets financed with funds from commercial banks.	129(56.3)
New investments financed by state-owned banks	Percentage of investments in new fixed assets financed with funds from state banks.	129(56.3)
New investments financed by non-banks financial institutions	Percentage of investments in new fixed assets financed with funds from non-banks financial institutions.	129(56.3)
New investments financed by trade credit	Percentage of investments in new fixed assets financed with credits from suppliers or customers.	129(56.3)
New investments financed by family/friends	Percentage of investments in new fixed assets financed with family/friends funds.	129(56.3)
New investments financed by informal sources	Percentage of investments in new fixed assets financed with funds from informal sources.	129(56.3)
Dummy for overdraft	Dummy that takes value 1 if the firm has access to an overdraft facility.	226(98.7)
Dummy for loan	Dummy that takes value 1 if the firm has access to a loan line.	220(96.1)
Rent land	Percentage of the land occupied for the establishment rented	229(100)
Rent Buildings	Percentage of the buildings occupied for the establishment rented	222(96.9)
Dummy for acquired or attempted to acquire additional land	Dummy that takes value 1 if the firm has acquired or attempted to acquire additional land.	223(97.4)
Dummy for purchased fixed assets	Dummy that takes value 1 if the firm has purchased fixed assets during last year.	227(99.1)

Table A.IV: Definitions of Investment climate (IC) variables of Quality, innovation and labor skills group.

Name of the variable	Description of the variable	Observations (Response rate %)
Dummy for quality certification	Dummy taking value 1 if the firm has any kind of quality certification.	225(98.3)
Dummy for new product	Dummy that takes value 1 if the plant has developed a new product line.	212(92.6)
Dummy for R&D	Dummy that takes value 1 if the firm performed R&D activities during last year.	217(94.8)
R&D expenditures	Total R&D expenditures as a percentage of sales	64(27.9)
Dummy for advertising	Dummy that takes value 1 if the firm has advertising expenditures.	226(98.7)
Advertising expenditures	Total advertising expenditures as a percentage of sales	124(54.1)
Dummy for payments for royalties	Dummy taking value 1 if the firm has paid royalties.	218(95.2)
Dummy for royalty payment received	Dummy taking value 1 if the firm has received royalties' payments.	117(51.1)
Annual royalty payments received	Total royalties payments received as a percentage of sales.	134(58.5)
Dummy for filed for any patents	Dummy taking value 1 if the firm has filed for any patents.	81(35.4)
Dummy for awarded any patents	Dummy taking value 1 if the firm has awarded any patents.	78(34.1)
Dummy for formally registered any copyrights	Dummy taking value 1 if the firm has formally registered any copyrights.	80(34.9)
Staff for which English is critical	Percentage of workforce for which English is critical.	222(96.9)
Dummy for internal training	Dummy taking value one if the firm provides formal (beyond on the job) internal training to its employees.	229(100)
Dummy for external training	Dummy taking value one if the firm provides formal (beyond on the job) external training to its employees.	229(100)

Table A.V: Definitions of Investment climate (IC) variables of Other control variables group.

Name of the variable	Description of the variable	Observations (Response rate %)
Age of the firm	Age of the firm in 2006.	229(100)
Dummy for incorporated company	Dummy that takes value 1 if the firm is an incorporated company.	220(96.1)
Dummy for FDI	Dummy that takes value 1 if any part of firm's capital is foreign.	223(97.4)
Dummy for importer	Dummy taking value 1 if the firm imports more than 10% of the total purchases of intermediate materials.	224(97.8)
Dummy for exporter	Dummy taking value 1 if the firm exports more than 10% of the total annual sales.	224(97.8)
Share of exports	Share of exports over total annual sales.	221(96.5)
Dummy for Software Technology Park	Dummy that takes value 1 if the firm is located in a Software Technology Park.	225(98.3)
Dummy for Export Processing Zone	Dummy that takes value 1 if the firm is located in a Export Processing Zone.	229(100)

Appendix B: Tables and figures

Table B.I: Total number of observations before and after cleaning missing values and outliers.

	Observations before cleaning	Observations after cleaning
Total number of firms surveyed	**359**	
Missing observations	**258**	**113**
of which:		
firms with one PF variable missing	90	2
firms with two PF variables missing	26	0
firms with three PF variables missing	31	0
firms with four PF variables missing	111	111
Outliers	**9**	**17**
of which:		
outliers in materials	5	4
outliers in labor cost	1	8
outliers in both materials and labor cost	3	5
Useful observations (outliers and missing excluded)	**92**	**229**

The cleaning process is performed in three steps*.

IV. Those firms with missing values in all the PF variables (sales, materials, labor cost and capital) are dropped from the sample. For the rest of the missing values we apply the procedure described in II and III.

V. We replace those observations with ratios materials to sales or labor cost to sales greater than one (outliers) following step III.

VI. We replace the missing values of the PF variables by their corresponding industry-region-size medians. If we do not have enough observations in some cells, we replace them by the corresponding industry-size medians. If we still do not have enough observations in those cells, in the next step we replace the missing values by the region-size medians. If still necessary, in the last step we compute the medians only by size and/or by industry to replace those missing values.

The last row of the table summarizes the number of useful observations for regression analysis before and after the cleaning process.

*See Escribano and Pena (2010),

Table B.II: Representativeness of production function variables before and after cleaning missing values and outliers.

a) by industry and state

		andhra pr		delhi		harayana		karnataka		maharasht		tamil nadu		uttar para		Total	
Industry		#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.
Telecom services	Before cleaning	0	0.0	1	0.3	3	0.8	2	0.6	4	1.1	3	0.8	0	0.0	13	3.6
	After cleaning	0	0.0	0	0.0	2	0.9	0	0.0	2	0.9	2	0.9	0	0.0	6	2.6
Software	Before cleaning	16	4.5	40	11.1	15	4.2	25	7.0	53	14.8	59	16.4	5	1.4	213	59.3
	After cleaning	15	6.6	39	17.0	12	5.2	14	6.1	18	7.9	45	19.7	4	1.7	147	64.2
ITES	Before cleaning	1	0.3	5	1.4	10	2.8	4	1.1	26	7.2	9	2.5	5	1.4	60	16.7
	After cleaning	0	0.0	3	1.3	10	4.4	1	0.4	6	2.6	8	3.5	3	1.3	31	13.5
Media & Entert.	Before cleaning	0	0.0	1	0.3	0	0.0	0	0.0	71	19.8	1	0.3	0	0.0	73	20.3
	After cleaning	0	0.0	1	0.4	0	0.0	0	0.0	43	18.8	1	0.4	0	0.0	45	19.7
Total	Before cleaning	17	4.7	47	13.1	28	7.8	31	8.6	154	42.9	72	20.1	10	2.8	359	100.0
	After cleaning	15	6.6	43	18.8	24	10.5	15	6.6	69	30.1	56	24.5	7	3.1	229	100.0

Source: Authors' calculations with India ICS data.

b) by industry and size

		micro		small		medium		large		Total	
Industry		#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.
Telecom services	Before cleaning	1	0.3	2	0.6	1	0.3	8	2.3	12	3.4
	After cleaning	1	0.4	2	0.9	0	0.0	3	1.3	6	2.7
Software	Before cleaning	17	4.9	106	30.3	29	8.3	58	16.6	210	60.0
	After cleaning	13	5.8	67	29.6	24	10.6	41	18.1	145	64.2
ITES	Before cleaning	5	1.4	30	8.6	4	1.1	19	5.4	58	16.6
	After cleaning	2	0.9	14	6.2	4	1.8	10	4.4	30	13.3
Media & Entert.	Before cleaning	15	4.3	44	12.6	1	0.3	10	2.9	70	20.0
	After cleaning	8	3.5	33	14.6	1	0.4	3	1.3	45	19.9
Total	Before cleaning	38	10.9	182	52.0	35	10.0	95	27.1	350	100.0
	After cleaning	24	10.6	116	51.3	29	12.8	57	25.2	226	100.0

Source: Authors' calculations with India ICS data.

Table C.I: Robust IC elasticities and semi-elasticities with respect to productivity – OLS Estimation.

Blocks of ICA variables		Two steps		Single step estimation			
		Solow residual		Cobb-Douglas		Translog	
	Explanatory ICA variables	Restric.	Unrestric.	Restric.	Unrestric.	Restric.	Unrestric.
Infrastructures	Dummy for own generator (b)	0.534*	0.524*	0.488*	0.481*	0.498*	0.399
	Dummy for high speed internet connection (b)	0.678	0.679	0.931**	0.914**	0.985**	0.668
	Dummy for internet to deliver services (b)	0.832**	0.821**	0.753**	0.765**	0.792**	0.603*
	Wait for an electric supply (a)	-0.946***	-0.925***	-0.967***	-0.910***	-1.127***	0.665
Red Tape, informality and others	Dummy for conflicts in courts (b)	-0.895**	-0.946**	-0.808**	-0.822***	-0.681*	-0.673**
	Manager's time spent in bureaucratic issues (b)	-0.02	-0.02	-0.019	-0.016	-0.013	-0.023*
	Number of inspections (a)	0.01	0.044	0.128	0.13	-0.12	1.257**
Finance and corporate governance	Working capital financed by state-owned banks (b)	-0.007	-0.007	-0.011**	-0.011**	-0.011**	-0.008
	Dummy for overdraft (b)	0.184	0.182	0.291	0.312	0.29	0.216
Quality, innovation and labor skills	R&D expenditures (a)	0.018*	0.017	0.026**	0.025**	0.018	-0.025
	Dummy for external training (b)	0.315*	0.287*	0.285	0.267	0.303	0.281
Other control variables	Dummy for incorporated company (b)	0.768*	0.788*	0.551	0.629	0.563	0.625
	Observations	226	226	226	226	226	226
	R-squared	0.293	0.297	0.756	0.759	0.768	0.792

NOTES:
Two steps estimation: in the first step estimation of equation (b2.1) by non-parametric techniques to compute productivity (Solow residual), in the second step estimate (3.2) and (3.3) by OLS using as dependent variable the Solow residual from the first step, either restricted or unrestricted.
Single step estimation: estimate (3.1), (3.2) and (3.3) in a single step by OLS, where (3.1) can be a Cobb-Douglas Production function or a Translogarithmic.
Restricted: equal input output for all the establishments in the country.
Unrestricted: equal input-output elasticities for all the establishments in the same sector.
*significant at 10%; ** significant at 5%; *** significant at 1% given by robust standard errors corrected for correlation between cluster (industry and region).
Each regression includes a set of industry, size and region dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.II: IC percentage contributions to aggregate log-productivity.

			Aggregate log-TFP	Average log-TFP	Allocative efficiency
Demean log-productivity	**Infrastructures**	Dummy for own generator (b)	8.39	4.35	4.04
		Dummy for high speed internet connection (b)	10.77	10.25	0.52
		Dummy for internet to deliver services (b)	13.15	11.88	1.27
		Wait for an electric supply (a)	-31.67	-23.40	-8.27
	Red Tape, informality and others	Dummy for conflicts in courts (b)	-2.36	-1.26	-1.10
		Manager's time spent in bureaucratic issues (b)	-0.50	-2.58	2.08
		Number of inspections (a)	0.10	0.13	-0.03
	Finance and Corporate Governance.	Working capital financed by state-owned banks (b)	-0.24	-0.50	0.26
		Dummy for overdraft (b)	2.64	0.97	1.67
	Quality, Innovation and labor skills	R&D expenditures (a)	2.46	2.26	0.21
		Dummy for external training (b)	4.83	2.44	2.39
	Other control variables	Dummy for incorporated company (b)	12.18	9.02	3.15
Total contribution of IC (demean log-productivity)			19.75	13.55	6.20
Other stuff		Industry/region/size controls	-20.00	-27.67	7.68
		Constant term	46.25	46.25	0.00
		Residual	53.99	0.00	53.99
Total contribution of other stuff			80.25	18.58	61.67
Total			100.00	32.13	67.87

NOTES:
Results from equation (5.3).
Demeaned log-productivity is the part of productivity associated with the investment climate
The productivity measure used is the restricted Solow residual.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.III: IC percentage contributions to aggregate log-productivity. (Exporter firms)

			Aggregate log-TFP	Average log-TFP	Allocative efficiency
Demean log-productivity	Infrastructures	Dummy for own generator (b)	11.77	8.88	2.89
		Dummy for high speed internet connection (b)	39.80	39.44	0.36
		Dummy for internet to deliver services (b)	25.80	23.45	2.35
		Wait for an electric supply (a)	-35.00	-33.96	-1.04
	Red Tape, informality and others	Dummy for conflicts in courts (b)	-5.21	-0.99	-4.22
		Manager's time spent in bureaucratic issues (b)	-0.59	-1.02	0.43
		Number of inspections (a)	13.50	9.67	3.83
	Finance and corporate governance.	Working capital financed by state-owned banks (b)	-1.65	-1.10	-0.55
		Dummy for overdraft (b)	3.84	1.93	1.91
	Quality, innovation and labor skills	R&D expenditures (a)	17.13	5.55	11.59
		Dummy for external training (b)	8.09	5.33	2.76
	Other control variables	Dummy for incorporated company (b)	-4.88	-4.62	-0.26
Total contribution of IC (demean log-productivity)			72.59	52.56	20.04
Other stuff		Industry/region/size controls	-44.59	-43.68	-0.90
		Constant term	61.23	61.23	0.00
		Residual	10.77	0.00	10.77
Total contribution of other stuff			27.41	17.54	9.87
Total			100.00	70.10	29.90

NOTES:
Results from equation (5.3).
Demeaned log-productivity is the part of productivity associated with the investment climate
The productivity measure used is the restricted Solow residual.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.IV: IC percentage contributions to aggregate log-productivity. (Non-exporter firms)

			Aggregate log-TFP	Average log-TFP	Allocative efficiency
Demean log-productivity	**Infrastructures**	Dummy for own generator (b)	7.80	2.69	5.11
		Dummy for high speed internet connection (b)	12.10	11.09	1.00
		Dummy for internet to deliver services (b)	23.20	22.75	0.45
		Wait for an electric supply (a)	-32.68	-16.98	-15.70
	Red Tape, informality and others	Dummy for conflicts in courts (b)	-1.00	-2.47	1.46
		Manager's time spent in bureaucratic issues (b)	-3.55	-8.11	4.56
		Number of inspections (a)	-23.29	-13.69	-9.60
	Finance and corporate governance	Working capital financed by state-owned banks (b)	-0.12	-0.18	0.06
		Dummy for overdraft (b)	-1.32	-1.26	-0.06
	Quality, innovation and labor skills	R&D expenditures (a)	12.74	3.41	9.33
		Dummy for external training (b)	3.79	2.20	1.59
	Other control variables	Dummy for incorporated company (b)	18.60	10.74	7.85
Total contribution of IC (demean log-productivity)			16.26	10.21	6.05
Other stuff		Industry/region/size controls	-17.50	-45.78	28.27
		Constant term	72.47	72.47	0.00
		Residual	28.77	0.00	28.77
Total contribution of other stuff			83.74	26.69	57.04
Total			100.00	36.91	63.09

NOTES:
Results from equation (5.3).
Demeaned log-productivity is the part of productivity associated with the investment climate
The productivity measure used is the restricted Solow residual.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table D.I: Robust IC elasticities and semi-elasticities with respect to productivity – OLS Estimation. (Software and ITES sub-sector)

Blocks of ICA variables	Explanatory ICA variables	Two steps		Single step estimation			
		Solow residual		Cobb-Douglas		Translog	
		Restric.	Unrestric.	Restric.	Unrestric.	Restric.	Unrestric.
Infrastructures	Dummy for own generator (b)	0.418*	0.417*	0.386*	0.436*	0.406*	0.305
	Dummy for high speed internet connection (b)	0.685	0.682	0.959*	1.002*	1.106*	0.7
	Dummy for unavailability of quality internet connection (b)	-0.800**	-0.776**	-0.781***	-0.777***	-0.766***	-0.687**
	Dummy for internet to deliver services (b)	0.860**	0.850**	0.804**	0.812**	0.765**	0.729**
Red Tape, informality and others	Dummy for conflicts in courts (b)	-0.682	-0.716	-0.57	-0.510*	-0.596*	-0.467
	Sales never repaid (a)	-0.022	-0.068	-0.119	-0.133	-0.056	-0.322**
	Manager's time spent in bureaucratic issues (b)	-0.014	-0.014	-0.013	-0.01	-0.005	-0.017
Finance and corporate governance	Working capital financed by state-owned banks (b)	-0.011	-0.011	-0.015**	-0.015**	-0.014**	-0.014**
	Dummy for overdraft (b)	0.386	0.382	0.507**	0.495**	0.508**	0.398
Quality, innovation and labor skills	Dummy for receiving royalty payments (b)	0.653*	0.65*	0.236	0.331	0.161	0.264
	Percentage of royalty payments received (b)	-0.927**	-0.927**	-0.831***	-0.850***	-0.795***	-0.855**
Other control variables	Dummy for incorporated company (b)	0.878**	0.874**	0.595	0.594	0.524	0.766*
	Observations	175	175	175	175	175	175
	R-squared	0.306	0.3	0.773	0.777	0.783	0.806

NOTES:
Two steps estimation: in the first step estimation of equation (b2.1) by non-parametric techniques to compute productivity (Solow residual), in the second step estimate (3.2) and (3.3) by OLS using as dependent variable the Solow residual from the first step, either restricted or unrestricted.
Single step estimation: estimate (3.1), (3.2) and (3.3) in a single step by OLS, where (3.1) can be a Cobb-Douglas Production function or a Translogarithmic.
Restricted: equal input output for all the establishments in the country.
Unrestricted: equal input-output elasticities for all the establishments in the same sector.
*significant at 10%; ** significant at 5%; *** significant at 1% given by robust standard errors corrected for correlation between cluster (industry and region).
Each regression includes a set of industry, size and region dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table D.II: IC percentage contributions to aggregate log-productivity. (Software and ITES sub-sector)

			Aggregate log-TFP	Average log-TFP	Allocative efficiency
Demean log-productivity	**Infrastructures**	Dummy for own generator (b)	6.478	4.068	2.410
		Dummy for high speed internet connection (b)	10.706	10.219	0.487
		Dummy for internet to deliver services (b)	-1.759	-2.000	0.241
		Wait for an electric supply (a)	13.360	11.900	1.460
	Red Tape, informality and others	Dummy for conflicts in courts (b)	-1.832	-0.974	-0.858
		Manager's time spent in bureaucratic issues (b)	-1.331	-1.044	-0.287
		Number of inspections (a)	-0.185	-1.617	1.433
	Finance and corporate governance	Working capital financed by state-owned banks (b)	-0.374	-0.831	0.457
		Dummy for overdraft (b)	5.657	2.415	3.243
	Quality, innovation and labor skills	Dummy for receiving royalty payments (b)	1.474	0.700	0.774
		Percentage of royalty payments received (b)	-0.004	-0.398	0.393
	Other control variables	Dummy for incorporated company (b)	13.701	11.994	1.708
Total contribution of IC (demean log-productivity)			45.893	34.433	11.460
Other stuff		Industry/region/size controls	0.342	0.829	-0.487
		Constant term	-0.872	-0.872	0.000
		Residual	54.637	0.000	54.637
Total contribution of other stuff			54.107	-0.043	54.150
Total			100.000	34.390	65.610

NOTES:
Results from equation (5.3).
Demeaned log-productivity is the part of productivity associated with the investment climate
The productivity measure used is the restricted Solow residual.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data

Table E.I.a: IC elasticities and semi-elasticities with respect to employment – IV Estimation.

Dependent variable: employment		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity1		0.339**	19.8	0.337**	19.1
Real wages1		-0.400**	-85.7	-0.398**	-85.3
Infrastructures	Dummy for own generator (b)	0.769***	9.4	0.771***	9.4
	Number of unavailability of quality internet connection (b)	-0.115**	-2.4	-0.117**	-2.5
Red Tape, informality and others	Dummy for conflicts in courts (b)	0.834**	2.1	0.850**	2.1
	Number of weeks to resolve overdue payments (a)	-0.569**	-11.9	-0.560**	-11.7
	Security cost (a)	-0.081**	-21.8	-0.084**	-22.5
Finance and corporate governance	Working capital financed by private commercial banks (b)	0.019***	2.9	0.019***	2.9
	Rent land (b)	-0.006***	-9.5	-0.006***	-9.6
	Dummy for acquiring additional land (b)	0.461*	2.7	0.452*	2.7
Quality, innovation and labor skills	Dummy for quality certification (b)	0.374*	3.3	0.381*	3.3
	Dummy for receiving royalty payments (b)	0.758*	0.9	0.754*	0.9
Other control variables	Age of the firm	0.626***	41.3	0.622***	41.1
	Percentage of direct exports (b)	0.011***	9.8	0.011***	9.6
	Dummy for exporting processing zone (b)	0.622**	1.3	0.611**	1.3
Instruments evaluation	First stage R-squared: productivity2	0.290		0.296	
	Partial R-squared: productivity3	0.096		0.095	
	Partial R-squared F test (p-value):productivity4	0.000		0.000	
	First stage R-squared: wages2	0.398		0.398	
	Partial R-squared: wages3	0.063		0.063	
	Partial R-squared F test (p-value): wages4	0.070		0.070	
	Hansen test (p-value)[5]	0.565		0.544	
	Observations	226		226	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity and real wages are endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with India ICS data.

Table E.I.b: IC elasticities and semi-elasticities with respect to employment – IV Estimation. (Software and ITES sub-sector)

Dependent variable: employment		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity		0.348*	21.5	0.343*	20.9
Real wages		-0.26	-56.9	-0.265	-57.9
Infrastructures	Dummy for own generator (b)	0.609***	8.8	0.619***	8.9
Red Tape, informality and others	Dummy for conflicts in courts (b)	0.758**	1.9	0.773**	2.0
	Number of weeks to resolve overdue payments (a)	-0.674***	-15.8	-0.663***	-15.6
	Manager's time spent in bureaucratic issues (b)	0.019***	4.2	0.019***	4.2
Finance and corporate governance	Working capital financed by private commercial banks (b)	0.012*	2.0	0.012*	2.0
	Dummy for loan (b)	0.625**	2.0	0.616**	2.0
	Rent land (b)	-0.009***	-16.1	-0.010***	-16.3
Quality, innovation and labor skills	Dummy for quality certification (b)	0.496**	4.7	0.497**	4.7
	Dummy for external training (b)	0.357*	4.8	0.368*	5.0
Other control variables	Age of the firm	0.583***	35.2	0.576***	34.8
	Percentage of direct exports (b)	0.009***	9.5	0.009***	9.5
	First stage R-squared: productivity	0.327		0.323	
	Partial R-squared: productivity	0.14		0.136	
	Partial R-squared F test (p-value):productivity	0.0002		0.0002	
	First stage R-squared: wages	0.345		0.345	
	Partial R-squared: wages	0.116		0.113	
	Partial R-squared F test (p-value): wages	0.0191		0.0191	
	Hansen test (p-value)5	0.194		0.193	
	Observations	175		175	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity and real wages are endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with India ICS data.

Table E.II.a: IC elasticities and semi-elasticities with respect to wages – IV Estimation

Dependent variable: real wages		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.444**	12.6	0.445**	12.3
Infrastructures	Dummy for unavailability of quality internet connection (b)	-0.694*	-1.5	-0.699*	-1.5
Red Tape, informality and others	Security cost expenditures (b)	-0.053	-5.0	-0.056*	-5.3
Finance and corporate governance	Working capital financed by family/friends (b)	-0.029**	-1.0	-0.029**	-1.0
	Dummy for acquired or attempted to acquire additional land (b)	0.448*	1.4	0.439*	1.4
Quality, innovation and labor skills	Dummy for R&D (b)	0.325	1.7	0.321	1.7
	Staff for which English is critical (b)	0.028***	17.1	0.028***	17.3
Other control variables	Percentage of direct exports (b)	0.007**	3.0	0.007**	3.0
Instruments evaluation	First stage R-squared[2]	0.264		0.268	
	Partial R-squared[3]	0.105		0.103	
	Partial R-squared F test (p-value)[4]	0.0001		0.0001	
	Hansen test (p-value)[5]	0.199		0.183	
	Observations	226		226	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with Pakistan ICS data.

Table E.II.b: IC elasticities and semi-elasticities with respect to wages – IV Estimation. (Software and ITES sub-sector)

Dependent variable: real wages		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.710***	20.7	0.710***	20.5
Infrastructures	Number of power outages (a)	-0.627*	-40.6	-0.615*	-39.8
Red Tape, informality and others	Sales never repaid (a)	-0.253	-11.4	-0.215	-9.7
Finance and corporate governance	Working capital financed by family/friends (b)	-0.038***	-1.2	-0.037***	-1.2
Quality, innovation and labor skills	Dummy for R&D (b)	0.357*	2.0	0.353*	2.0
Other control variables	Age of the firm	-0.459**	-13.2	-0.463**	-13.3
Instruments evaluation	First stage R-squared[2]	0.235		0.228	
	Partial R-squared[3]	0.09		0.09	
	Partial R-squared F test (p-value)[4]	0.019		0.019	
	Hansen test (p-value)[5]	0.845		0.87	
	Observations	175		175	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with Pakistan ICS data.

Table E.III.a: IC linear probability coefficients with respect to the probability of exporting – IV Estimation

Dependent variable: export		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.114*	55.4	0.115*	54.2
Infrastructures	Dummy for own generator (b)	0.145*	14.6	0.145*	14.6
	Dummy for accommodation constrain (b)	-0.166**	-36.9	-0.162**	-36.1
Red Tape, informality and others	Dummy for conflicts in courts (b)	0.187*	4.0	0.193*	4.2
	Sales never repaid (a)	-0.083*	-94.8	-0.080*	-91.5
Finance and corporate governance	Working capital financed by informal sources (b)	-0.026***	-2.2	-0.026***	-2.2
	New investments financed by new equity (a)	0.069***	3.9	0.069***	3.9
Quality, innovation and labor skills	Staff for which English is critical (b)	0.002*	24.8	0.002*	25.3
	Dummy for external training (a)	0.11*	13.5	0.114*	13.9
Other control variables	Age of the firm	-0.126***	-69.8	-0.127***	-70.3
	Dummy for Export Processing Zone (b)	0.182*	3.2	0.179*	3.2
Instruments evaluation	First stage R-squared[2]	0.333		0.341	
	Partial R-squared[3]	0.0596		0.0605	
	Partial R-squared F test (p-value)[4]	0.0103		0.0103	
	Hansen test (p-value)[5]	0.496		0.506	
	Observations	222		222	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with Pakistan ICS data.

Table E.III.b: IC linear probability coefficients with respect to the probability of exporting – IV Estimation. (Software and ITES sub-sector)

Dependent variable: export		**Restricted Solow residual**		**Unrestricted Solow residual**	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.265***	116.5	0.265***	115.0
Infrastructures	Dummy for internet to make purchases (b)	0.191**	21.9	0.191**	22.0
	Dummy for accommodation constrain (b)	-0.217**	-37.6	-0.209**	-36.1
Red Tape, informality and others	Dummy for conflicts in courts (b)	0.280*	5.3	0.294**	5.5
	Security cost (a)	-0.087***	-133.0	-0.088***	-135.2
Finance and corporate governance	Working capital financed by family/friends (b)	-0.003***	-2.1	-0.003***	-2.1
	Working capital financed by informal sources (b)	-0.030***	-2.3	-0.030***	-2.2
Quality, innovation and labor skills	Dummy for quality certification (b)	0.151	10.4	0.151	10.4
	Royalties received (b)	0.258**	0.9	0.259**	0.9
Other control variables	Age of the firm	-0.102	-43.9	-0.104	-44.8
Instruments evaluation	First stage R-squared[2]	0.291		0.287	
	Partial R-squared[3]	0.0635		0.0637	
	Partial R-squared F test (p-value)[4]	0.056		0.0548	
	Hansen test (p-value)[5]	0.667		0.667	
	Observations	171		171	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with Pakistan ICS data.

Table E.IV:a: IC linear probability coefficients with respect to the probability of receiving FDI – IV Estimation

Dependent variable: probability of receiving fdi		**Restricted Solow residual**		**Unrestricted Solow residual**	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.152**	199.7	0.147**	187.9
Infrastructures	Dummy for own transport for workers (b)	0.173*	11.4	0.169	11.2
Red Tape, informality and others	Dummy for cyber crime (b)	-0.418***	-7.7	-0.417***	-7.7
Finance and corporate governance	Working capital financed by informal sources (b)	-0.016**	-4.5	-0.015**	-4.4
	Rent land (b)	0.001**	50.7	0.001**	49.6
Quality, innovation and labor skills	Staff for which English is critical (b)	0.004**	108.4	0.004**	106.7
Other control variables	Share of exports (b)	0.003***	49.2	0.003***	48.7
Instruments evaluation	First stage R-squared[2]	0.263		0.268	
	Partial R-squared[3]	0.0379		0.377	
	Partial R-squared F test (p-value)[4]	0.0603		0.0601	
	Hansen test (p-value)[5]	0.578		0.554	
	Observations	217		217	

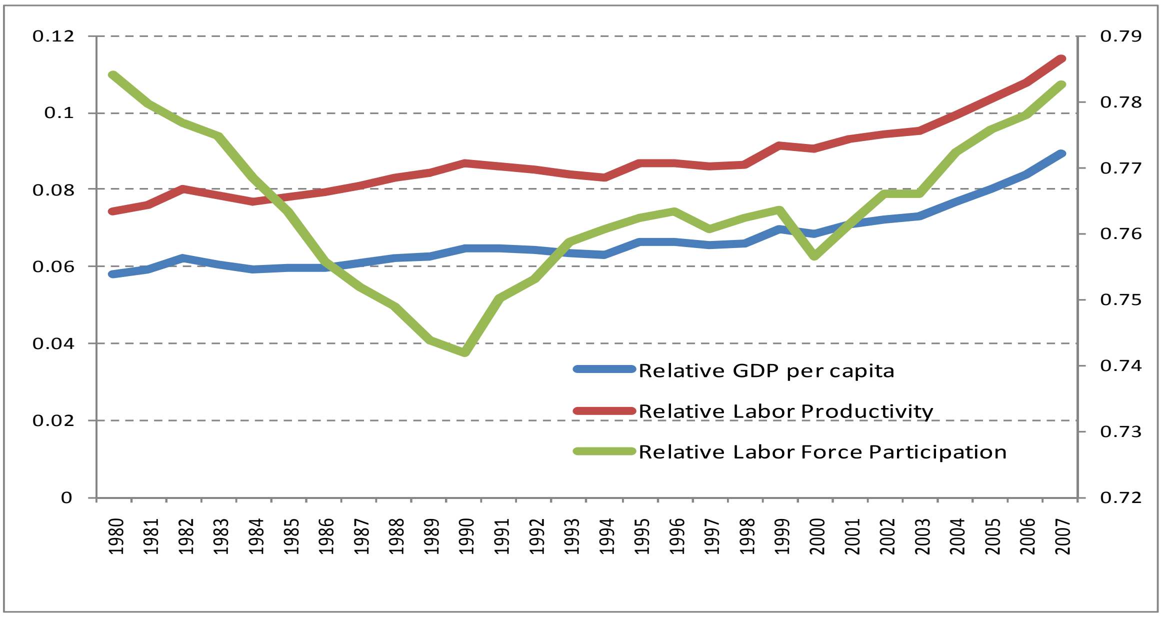
NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with Pakistan ICS data.

Table E.IV:b: IC linear probability coefficients with respect to the probability of receiving FDI – IV Estimation. (ITES and Software sub-sector)

Dependent variable: probability of receiving fdi		Restricted Solow residual		Unrestricted Solow residual	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Productivity[1]		0.189**	222.3	0.189**	219.0
Infrastructures	Dummy for own transport for workers(b)	-0.079*	-212.0	-0.080*	-215.4
	Number of power outages (b)	0.234*	16.4	0.231*	16.1
Red Tape, informality and others	Dummy for cyber crime (b)	-0.445***	-9.0	-0.446***	-9.0
Finance and corporate governance	Working capital financed by informal sources (b)	-0.020**	-4.7	-0.020**	-4.7
	Rent land (b)	0.002**	61.9	0.002**	61.3
Quality, innovation and labor skills	Staff for which English is critical (b)	0.004***	104.6	0.004***	103.9
Other control variables	Share of exports (b)	0.002**	44.4	0.002**	43.5
Instruments evaluation	First stage R-squared[2]	0.321		0.314	
	Partial R-squared[3]	0.0827		0.0831	
	Partial R-squared F test (p-value)[4]	0.0435		0.0399	
	Hansen test (p-value)[5]	0.992		0.996	
	Observations	166		166	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region).
Each regression includes a set of industry, region and size dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Productivity is endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
Source: Authors' calculations with Pakistan ICS data.

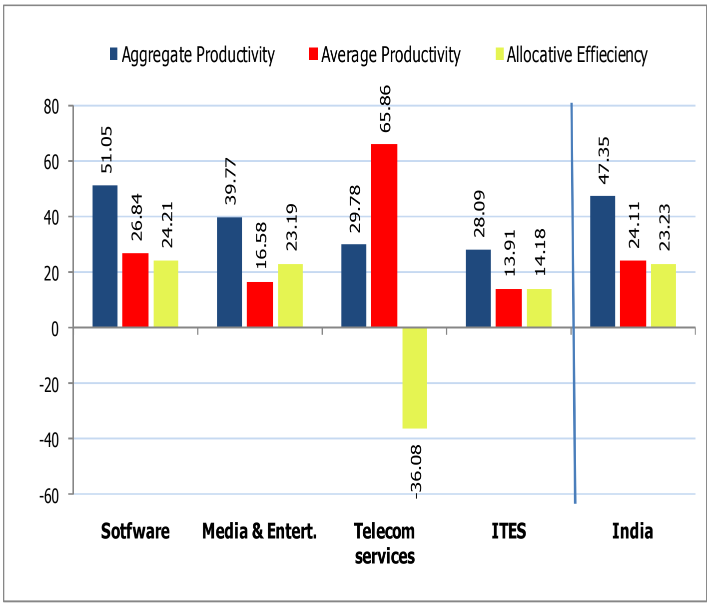
Figure 5: Decomposition of GDP gap between India and USA, 1980/2007



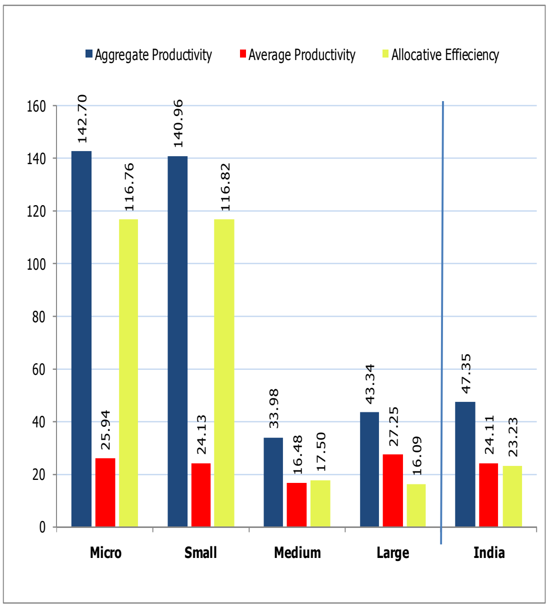
Source: Authors `Calculations with Penn World Table Version 6.3, Center for International Comparisons at the University of Pennsylvania.

Figure 6: Olley and Pakes decomposition in levels.

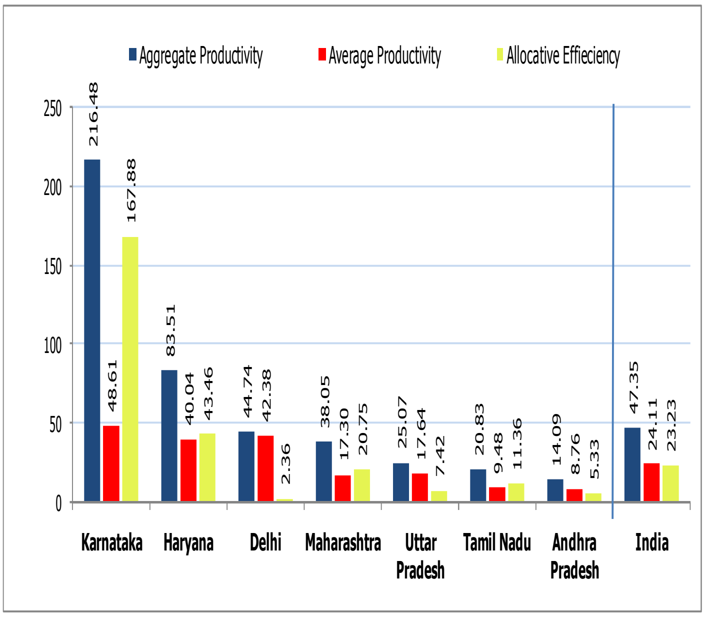
e) by Industry



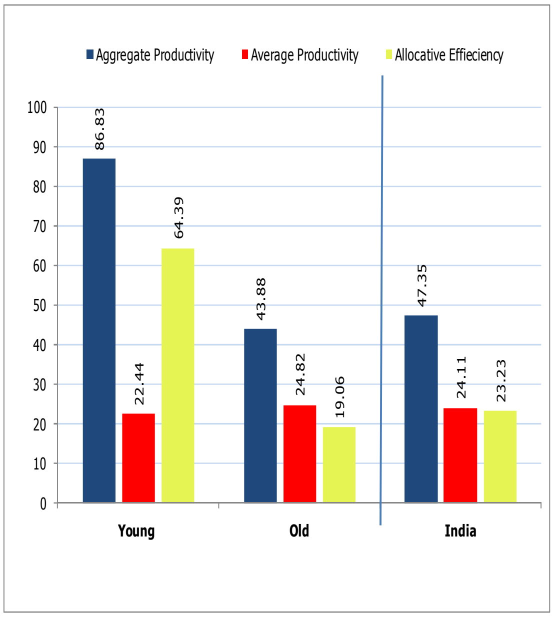
f) by Size



g) by State



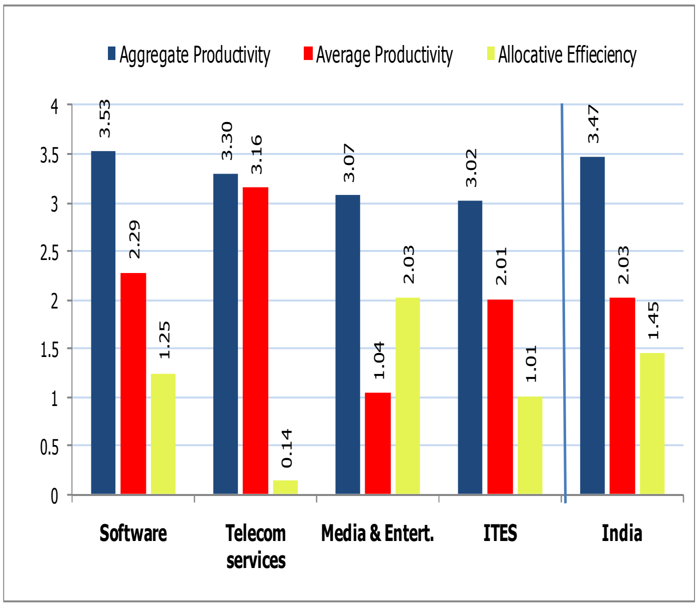
h) by Age



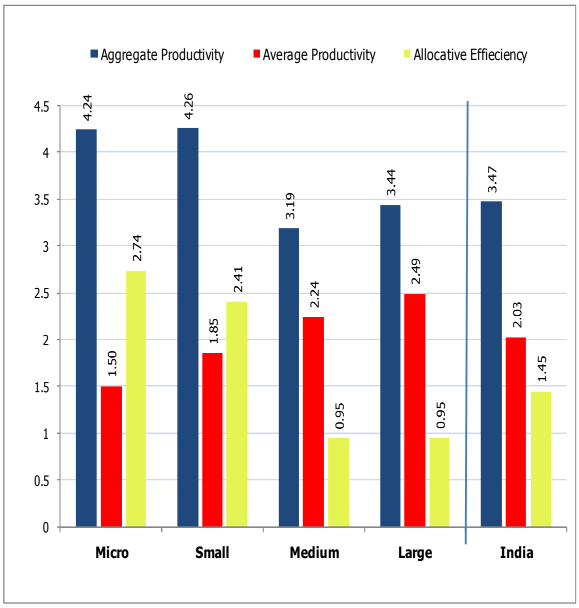
Note: Olley and Pakes decomposition in levels according to equation (5.1a). The productivity measure used is the restricted Solow residual in levels.
Source: Authors' calculations with India ICS data.

Figure 7: *Mixed* Olley and Pakes decomposition.

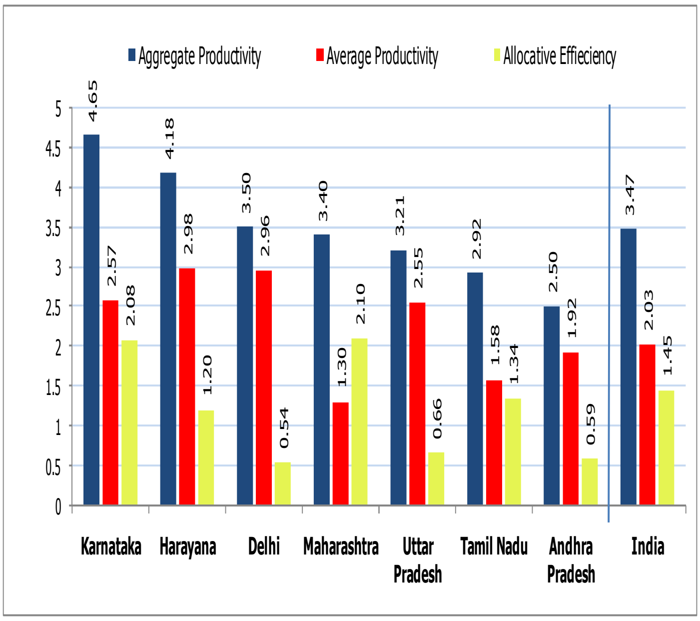
e) by Industry



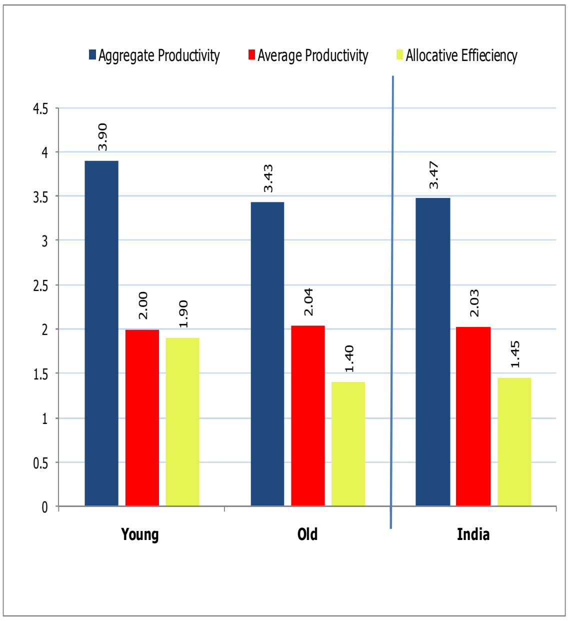
f) by Size



g) by State

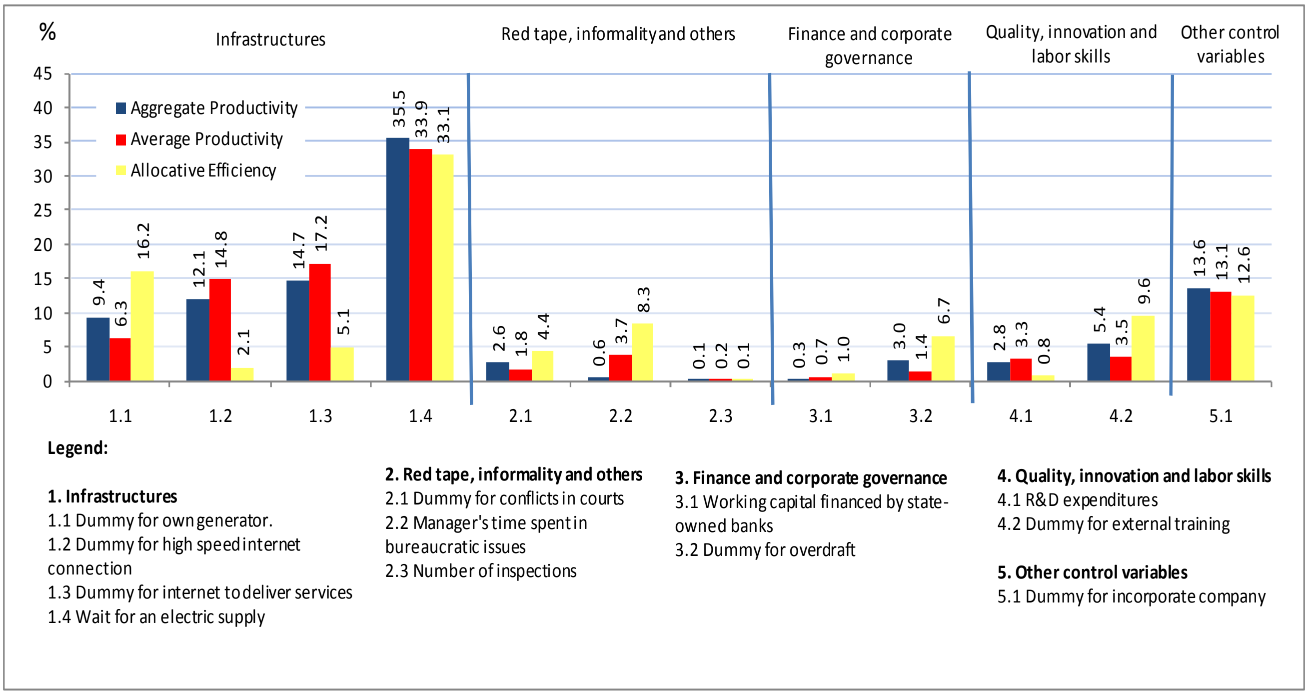


h) by Age



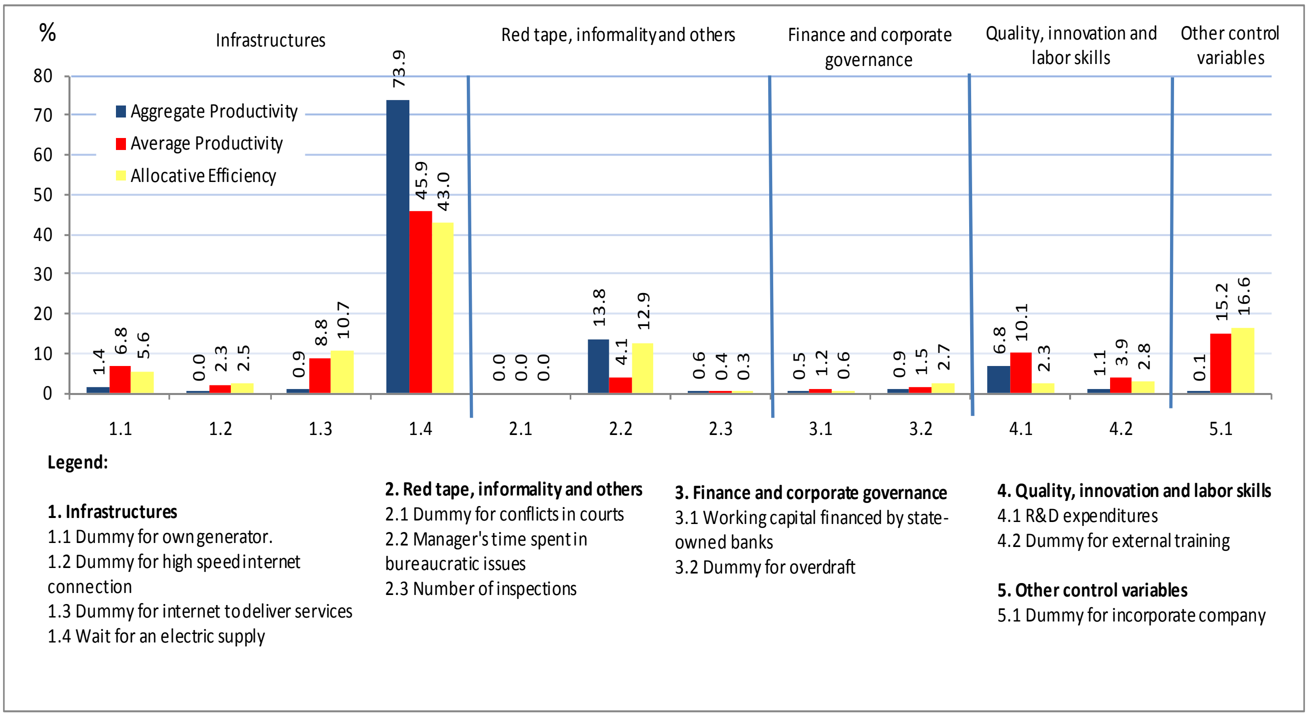
Note: Mixed Olley and Pakes decomposition according to equation (5.1b). The productivity measure used is the restricted Solow residual in logs.
Source: Authors' calculations with India ICS data.

Figure 4.1: IC Percentage contributions to aggregate log-productivity



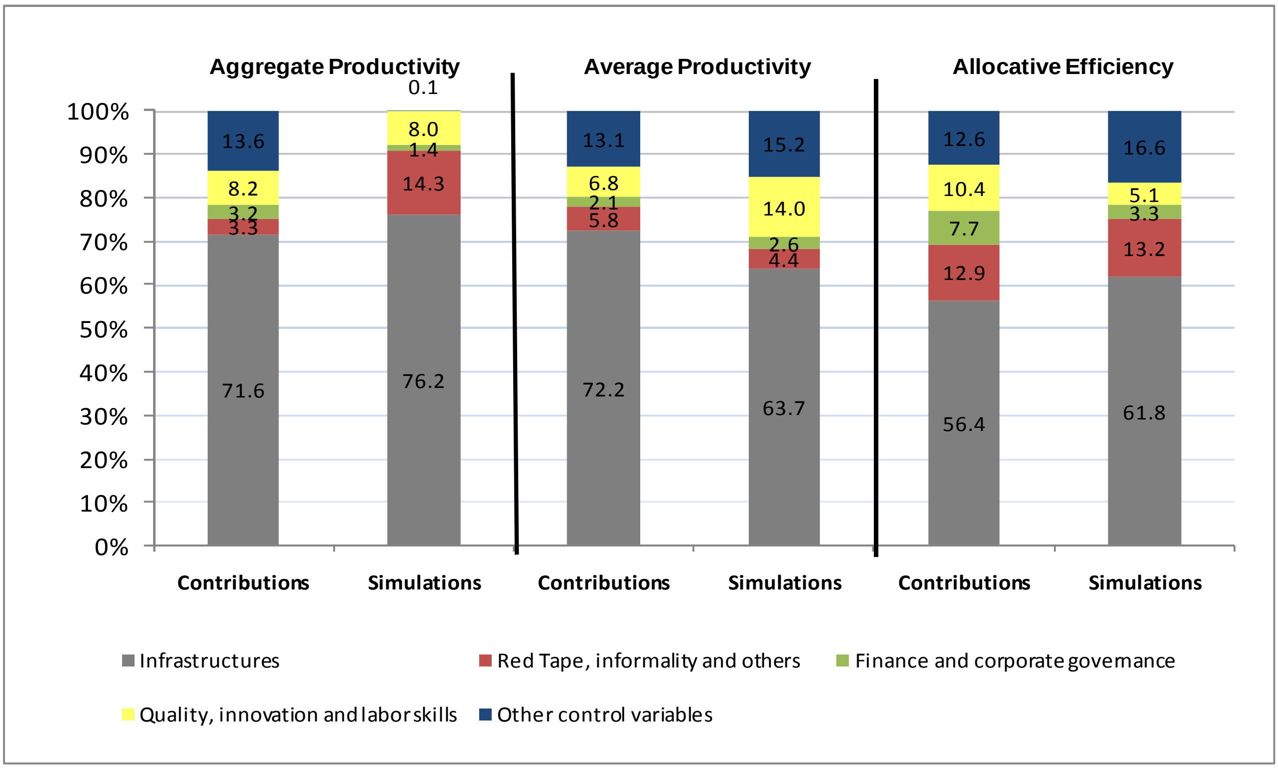
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 4.2: Percentage change in aggregate productivity (TFP) from a 20% improvement of IC variables.



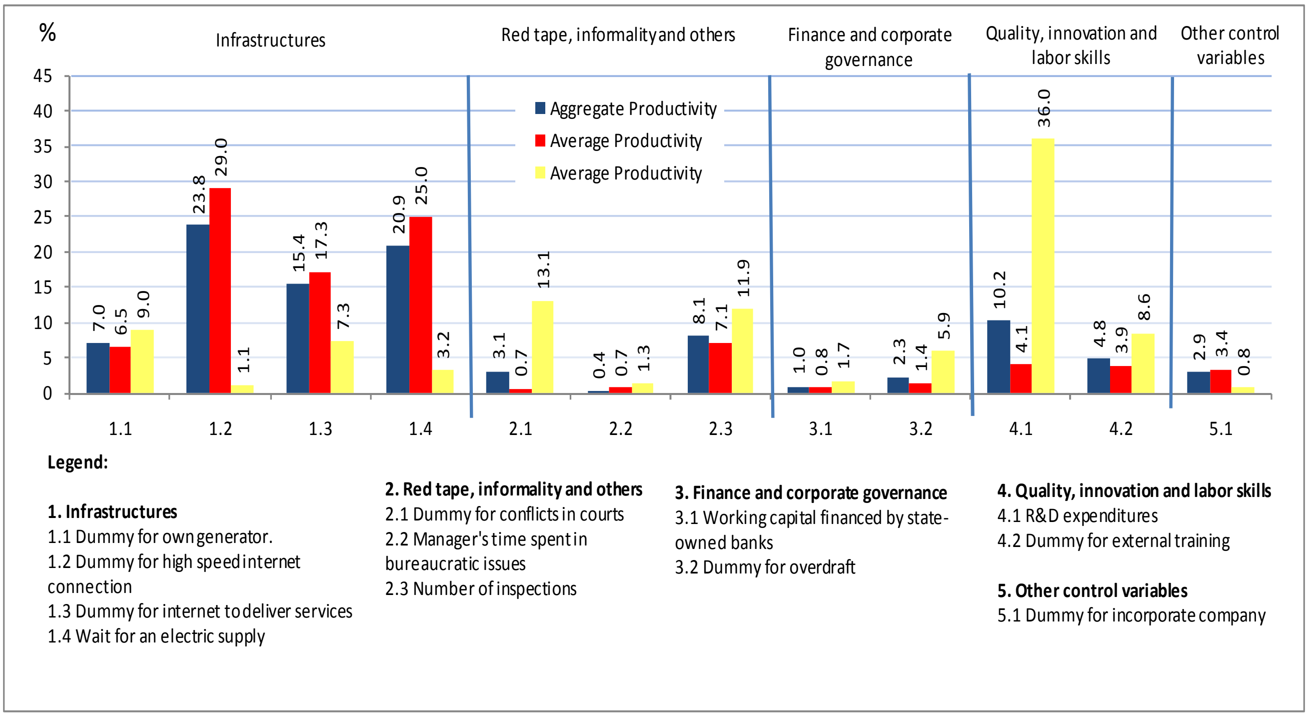
Note: Simulations computed according to section 4.3. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 4.3: Weight of each block of IC variables on aggregate productivity, average productivity and allocative efficiency, by contributions and by simulations.



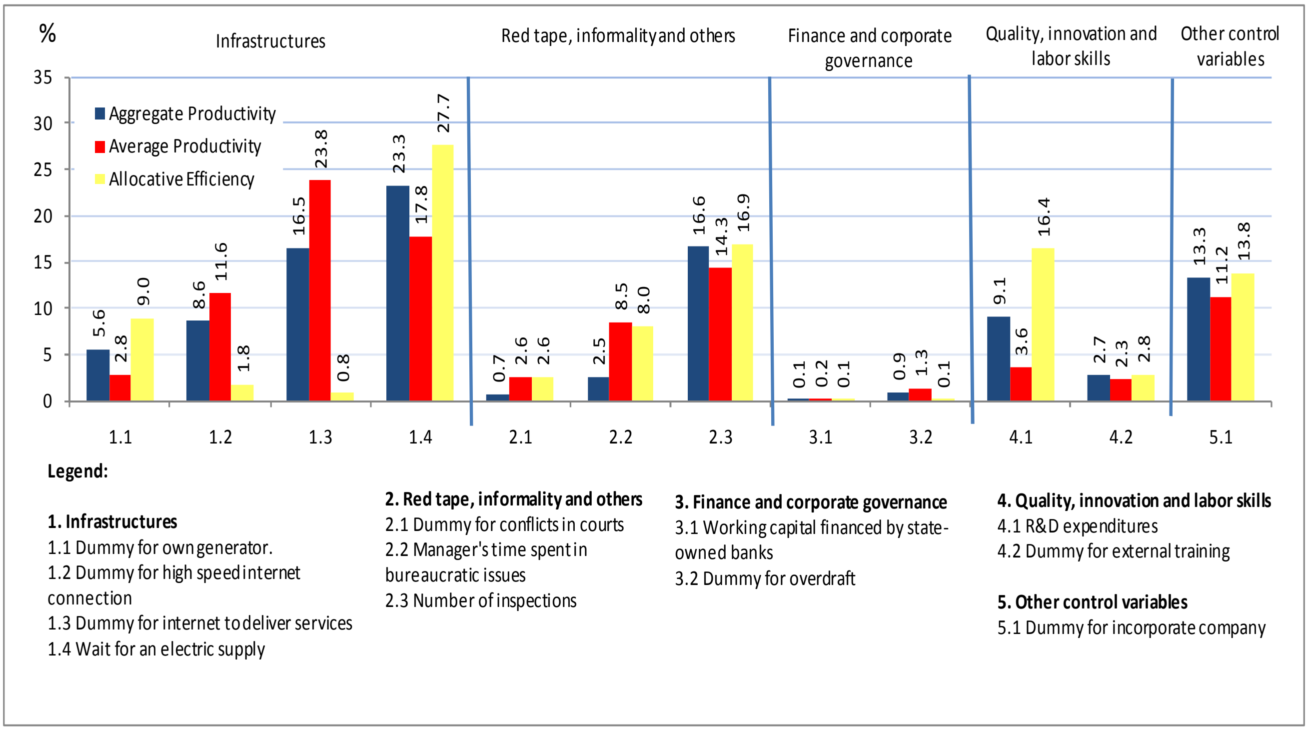
Note: The weight of each block or group of IC variables from contributions comes from Figure 4.1. We take the percentage contributions of Figure 4.1 in absolute value and we compute the relative weight of each block.
For the case of simulations we do the same with the percentage increases of Figure 4.2.
The productivity measure used is the *demeaned* restricted Solow residual in logs.
Source: Authors' calculations with India ICS data.

Figure 4.4: IC Percentage contributions to aggregate log-productivity. Exporters



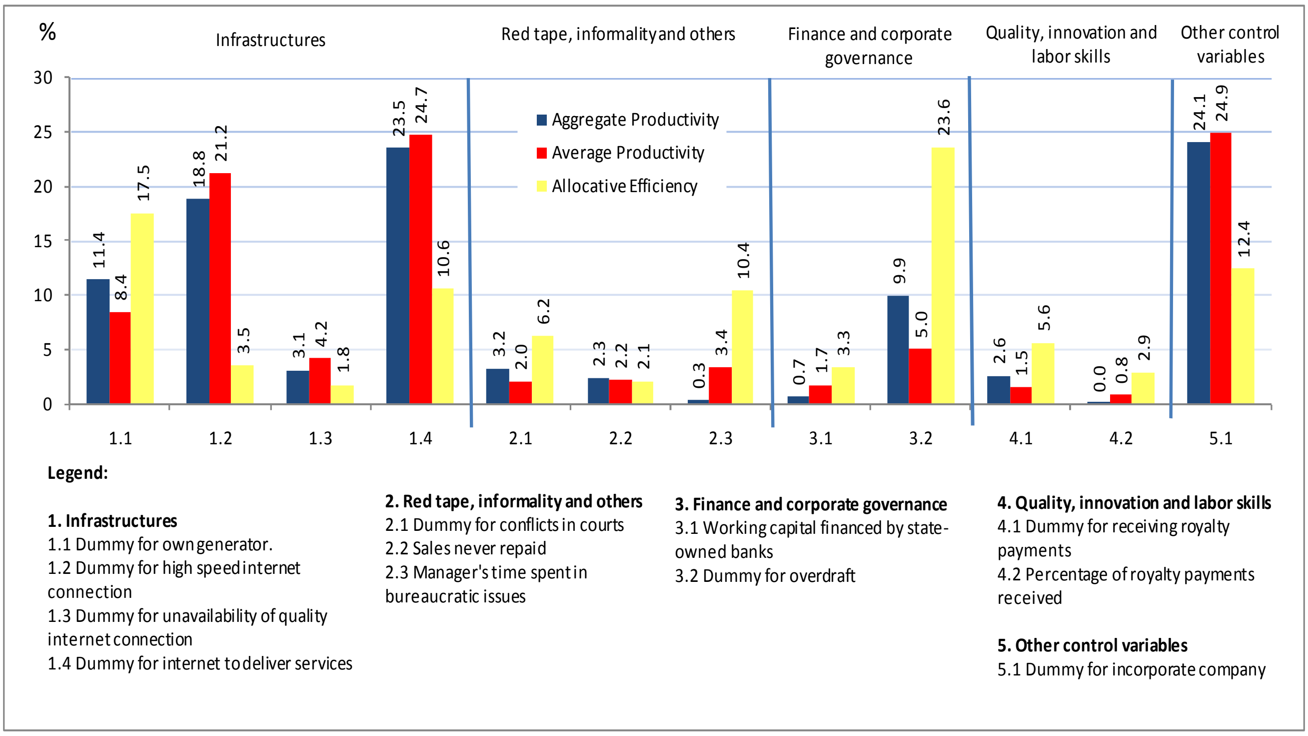
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 4.5: IC Percentage contributions to aggregate log-productivity. Non-Exporters



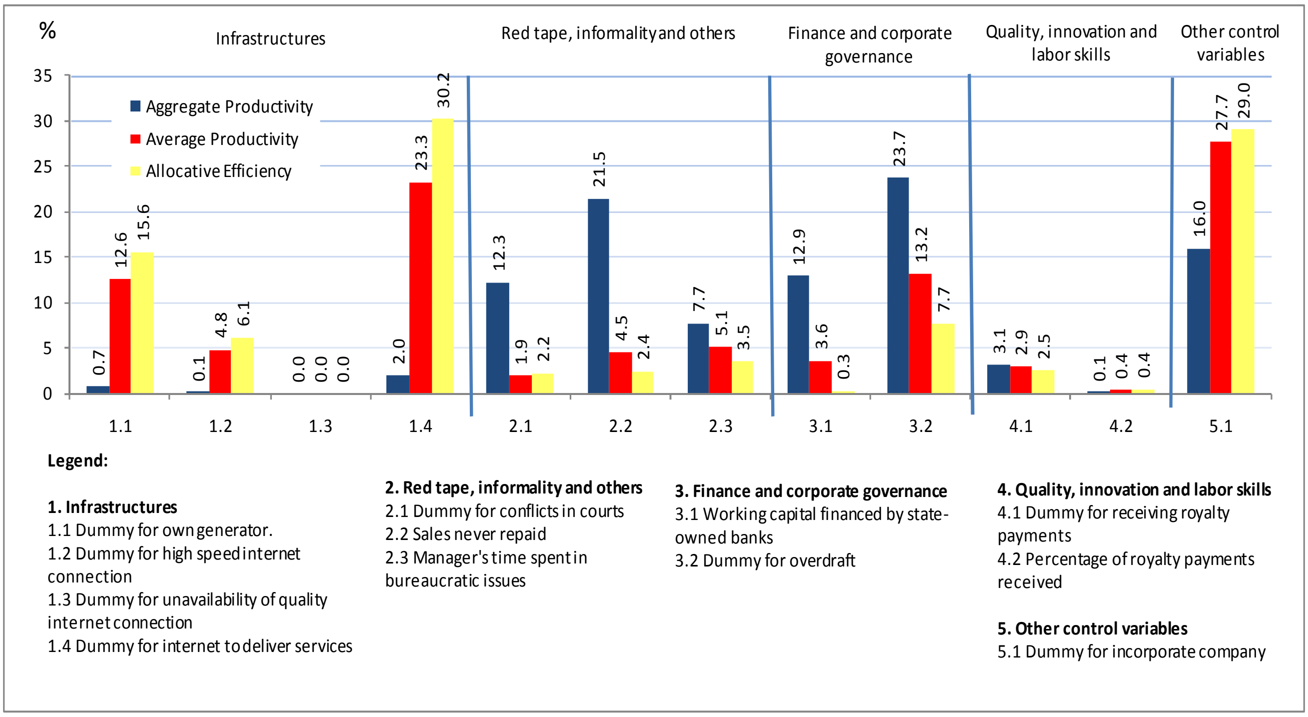
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 4.6: IC Percentage contributions to aggregate log-productivity. (ITES and Software sub-sector)



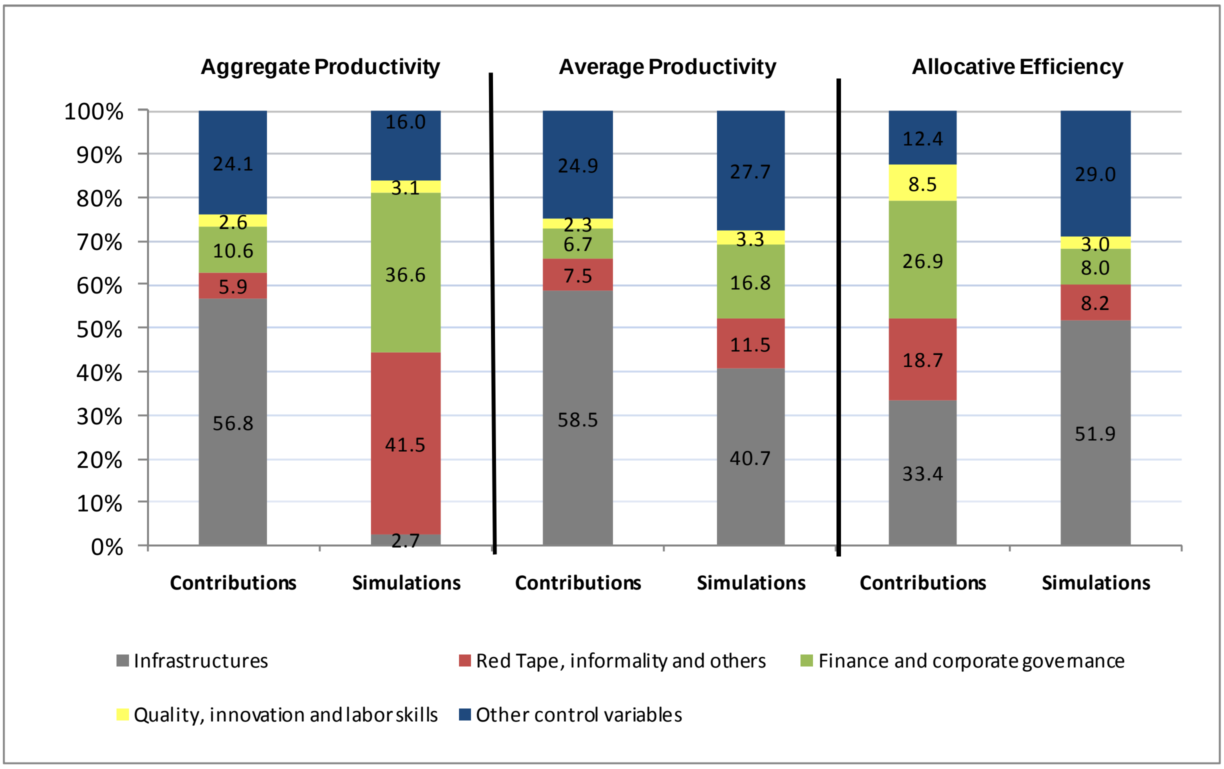
Note: Contributions computed according to section 5.4.The productivity measure used is the restricted Solow residual.
Source: Authors' calculations with India ICS data.

Figure 4.7: Percentage change in aggregate productivity (TFP) from a 20% improvement of IC variables. (ITES and Software sub-sector)



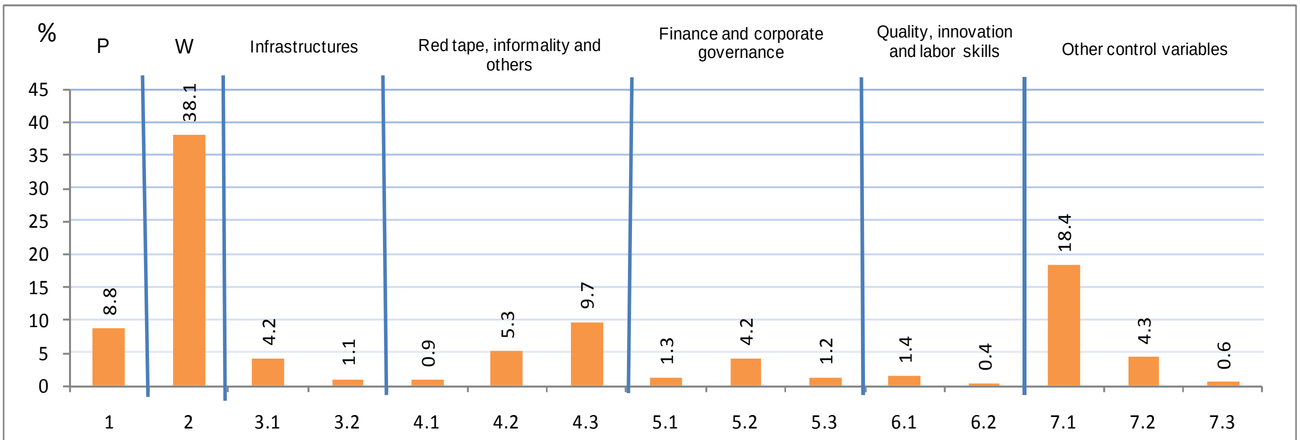
Note: Simulations computed according to section 4.3. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 4.8: Weight of each block of IC variables on aggregate productivity, average productivity and allocative efficiency, by contributions and by simulations. (ITES and Software sub-sector)



Note: The weight of each block or group of IC variables from contributions comes from Figure 4.6. We take the percentage contributions of Figure 4.6 in absolute value and we compute the relative weight of each block.
For the case of simulations we do the same with the percentage increases of Figure 4.7.
The productivity measure used is the *demeaned* restricted Solow residual in logs.
Source: Authors' calculations with India ICS data.

Figure 5.1: IC percentage contributions to average log-employment.



Legend:

1. Log of Productivity
2. Log of Real Wage

3. Infrastructures
3.1 Dummy for own generator
3.2 Number of unavailability of quality internet connection

4. Red tape, informality and others
4.1 Dummy for conflicts with court involved
4.2 Number of weeks to resolve overdue payments
4.3 Security cost

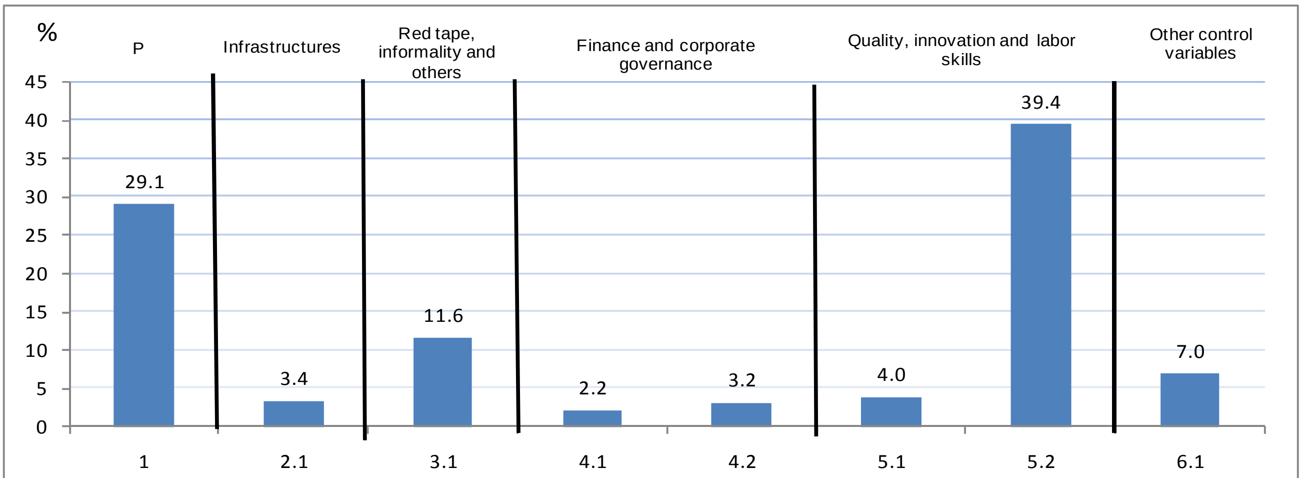
5. Finance and corporate governance
5.1 Working capital financed by private commercial banks
5.2 Rent land
5.3 Dummy for acquiring additional land

6. Quality, innovation and labor skills
6.1 Dummy for quality certification
6.2 Dummy for receiving royalty payments

7. Other control variables
7.1 Age of the firm
7.2 Percentage of direct exports

Note: Contributions computed according to section 5.5. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.2: IC percentage contributions to average log-wage.



Legend:

1. Log of Productivity

2. Infrastructures
2.1 Dummy for unavailability of quality internet connection

3. Red tape, informality and others
3.1 Security cost expenditures

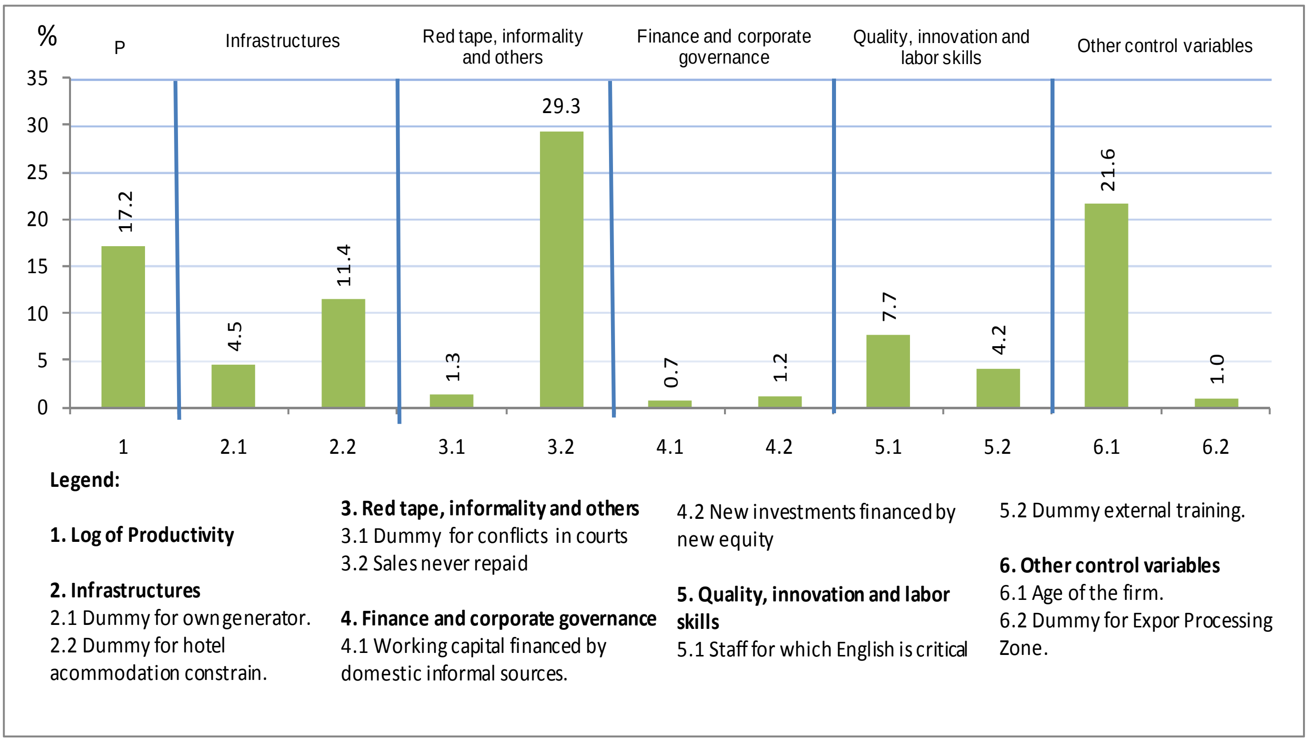
4. Finance and corporate governance
4.1 Working capital financed by family/friends
4.2 Dummy for acquired or attempted to acquire additional land

5. Quality, innovation and labor skills
5.1 Dummy for R&D.
5.2 Staff for English is critical

6. Other control variables
6.1 Percentage of direct exports

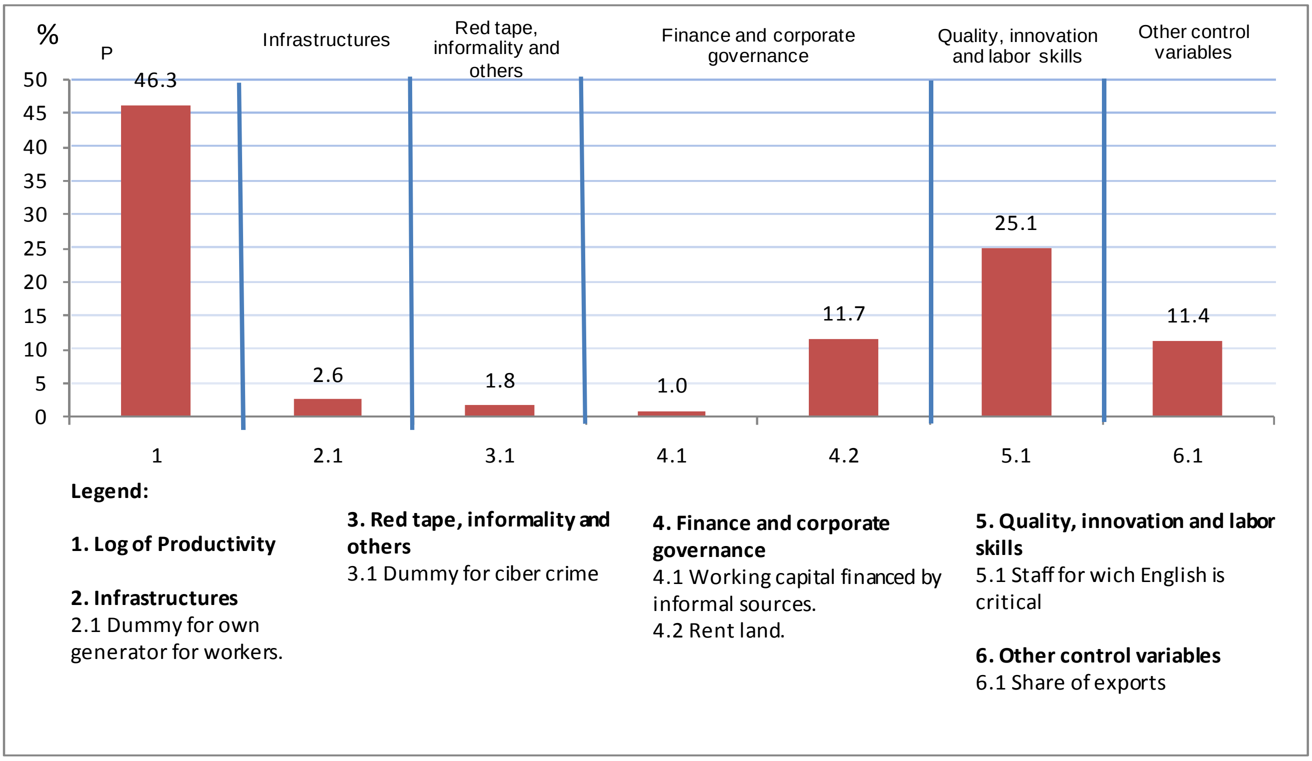
Note: Contributions computed according to section 5.6. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.3: IC percentage contributions to the probability of exporting.



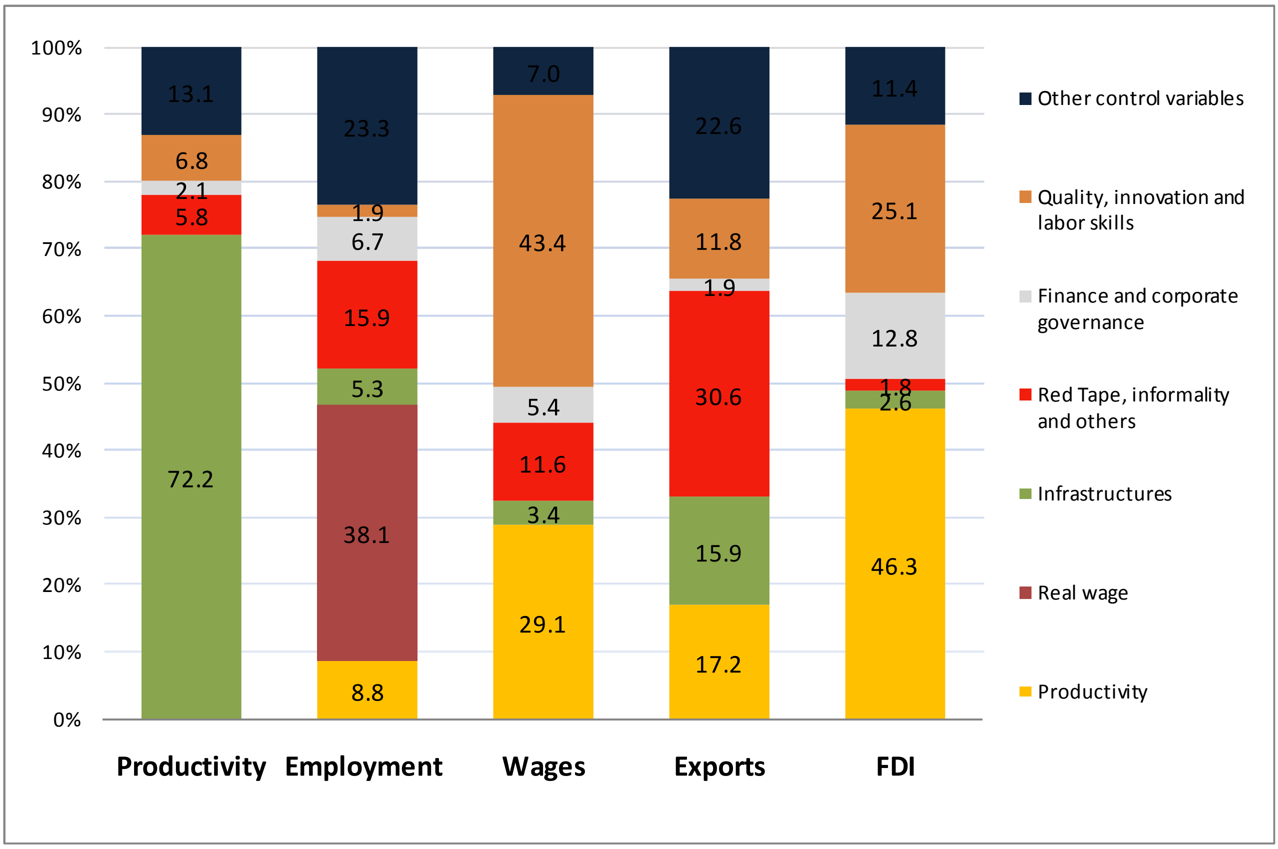
Note: Simulations computed according to section 5.7. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.4: IC percentage contributions to the probability of receiving FDI



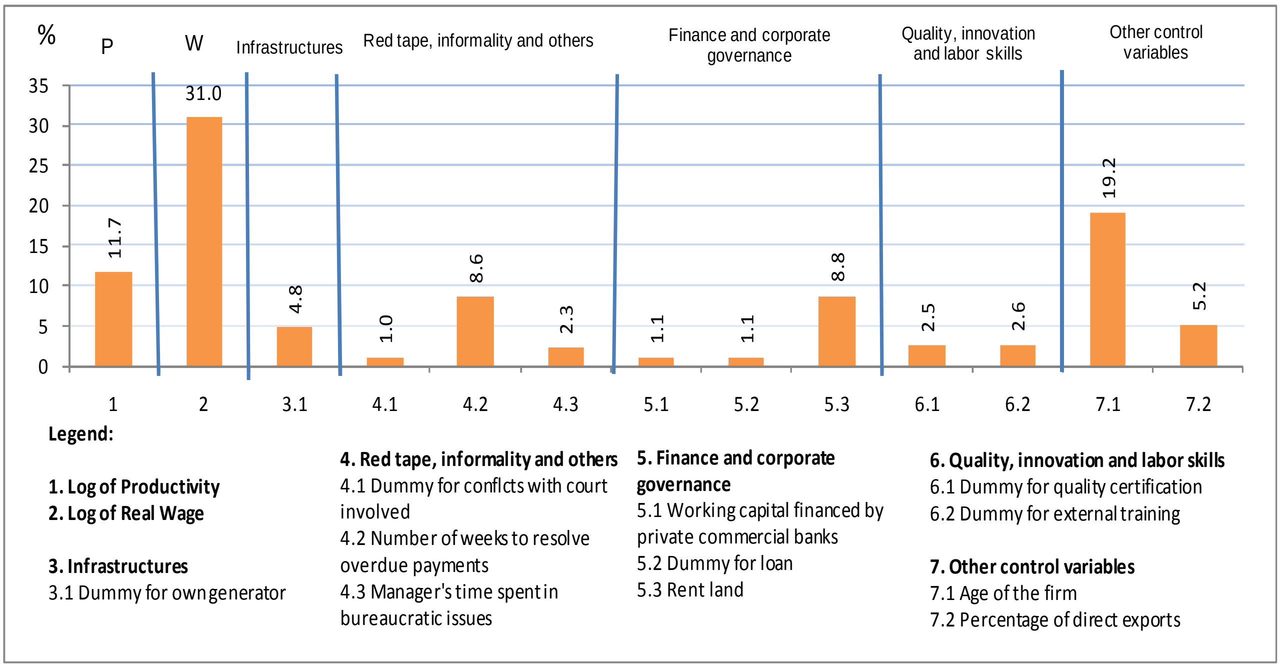
Note: Simulations computed according to section 5.8. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.5: Weight of each block of IC variables on the sample means of economic performance measures



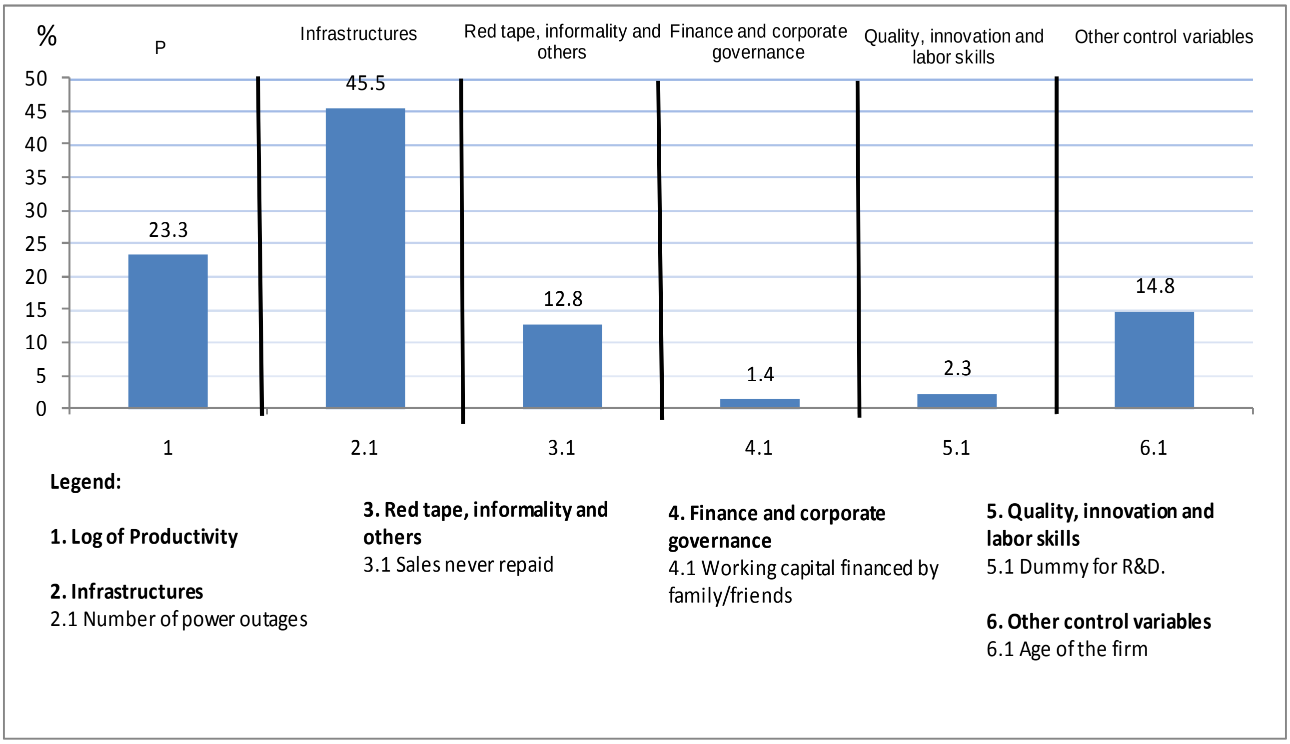
Note: The weight of each block or group of IC variables to the sample means comes from the contributions of Figures 5.1 to 5.4. We take the percentage contributions in absolute value and we compute the relative weight of each block. For the case of productivity we take the IC contributions to average log-productivity from Figure 4.1.
Source: Authors' calculations with India ICS data.

Figure 5.6: IC percentage contributions to average log-employment. (Software and ITES sub-sector)



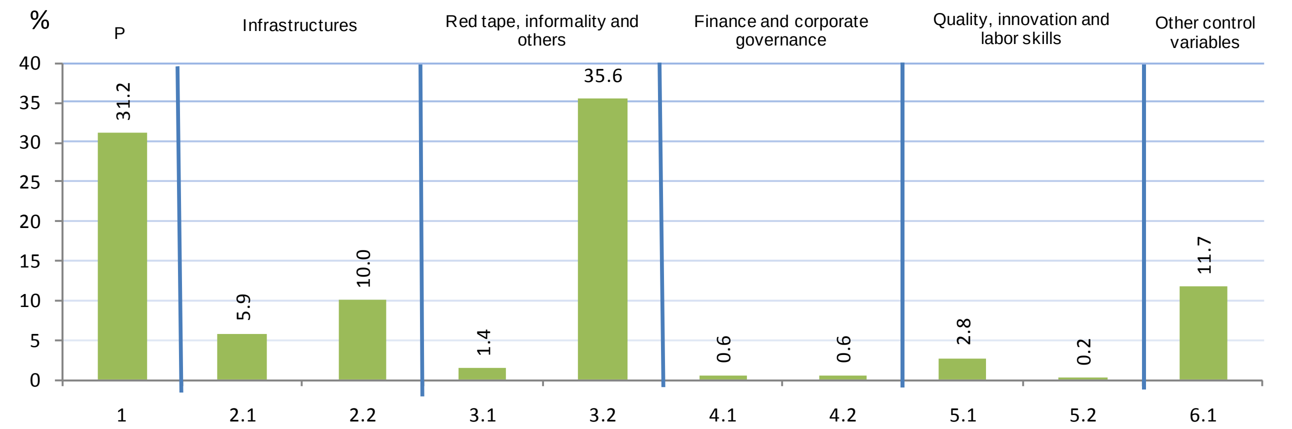
Note: Contributions computed according to section 5.5. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.7: IC percentage contributions to average log-wage. (Software and ITES sub-sector)



Note: Contributions computed according to section 5.6. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.8: IC percentage contributions to the probability of exporting. (Software and ITES sub-sector)



Legend:

1. Log of Productivity

2. Infrastructures
2.1 Dummy for internet to make purchases.
2.2 Dummy for hotel acommodation constrain.

3. Red tape, informality and others
3.1 Dummy for conflicts in court s.
3.2 Security cost.

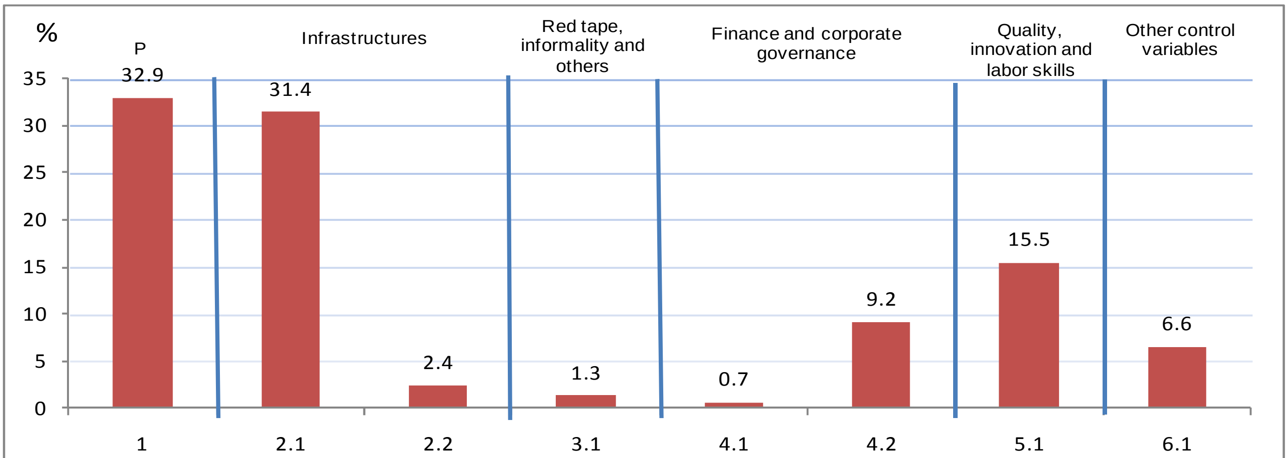
4. Finance and corporate governance
4.1 Working capital financed by domestic family/friends.
4.2 Working capital financed by domestic informal sources.

5. Quality, innovation and labor skills
5.1 Dummy for quality certification.
5.2 Royalties received.

6. Other control variables
6.1 Age of the firm.

Note: Contributions computed according to section 5.7. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.9: IC percentage contributions to the probability of receiving FDI. (Software and ITES sub-sector)



Legend:

1. Log of Productivity

2. Infrastructures
2.1 Dummy for own generator for workers.
2.2 Number of power outages

3. Red tape, informality and others
3.1 Dummy for ciber crime

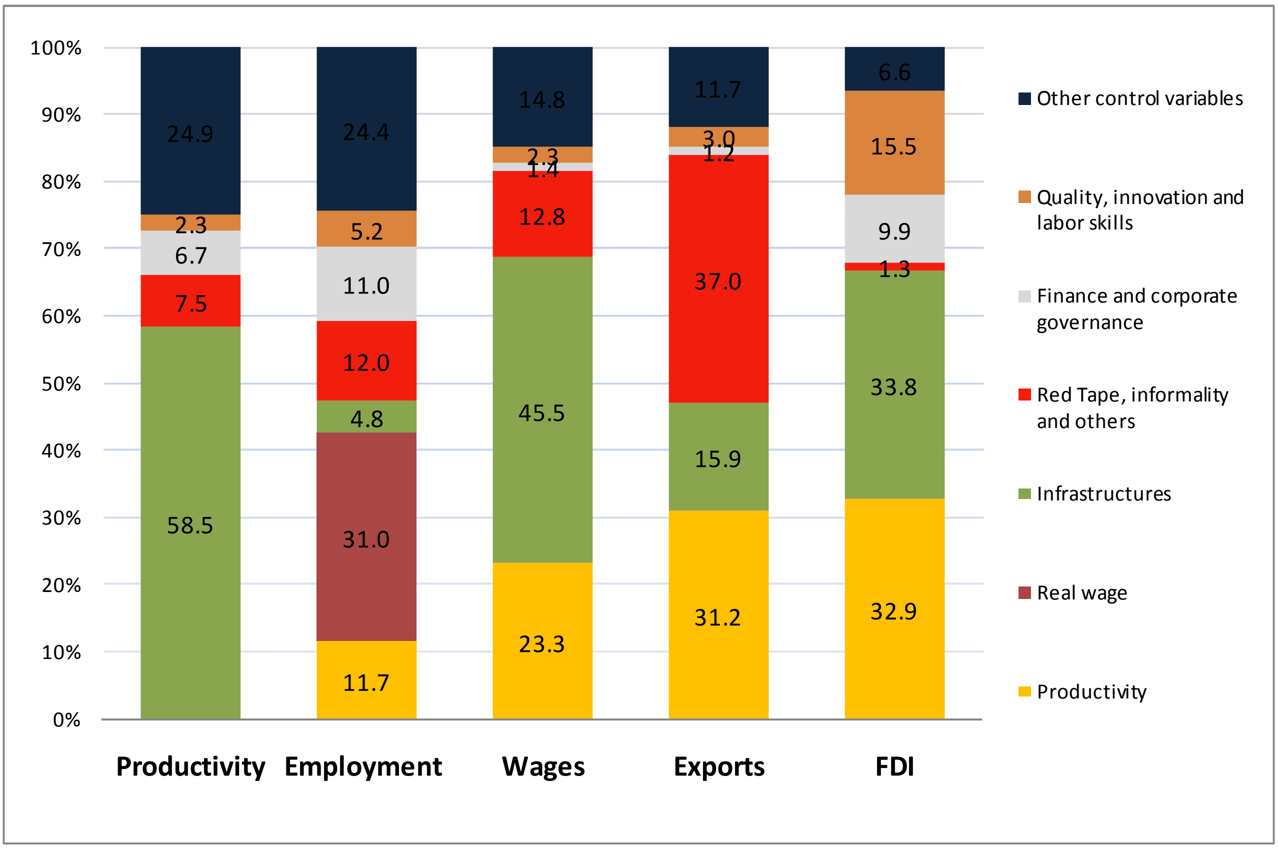
4. Finance and corporate governance
4.1 Working capital financed by informal sources.
4.2 Rent land.

5. Quality, innovation and labor skills
5.1 Staff for wich English is critical

6. Other control variables
6.1 Share of exports

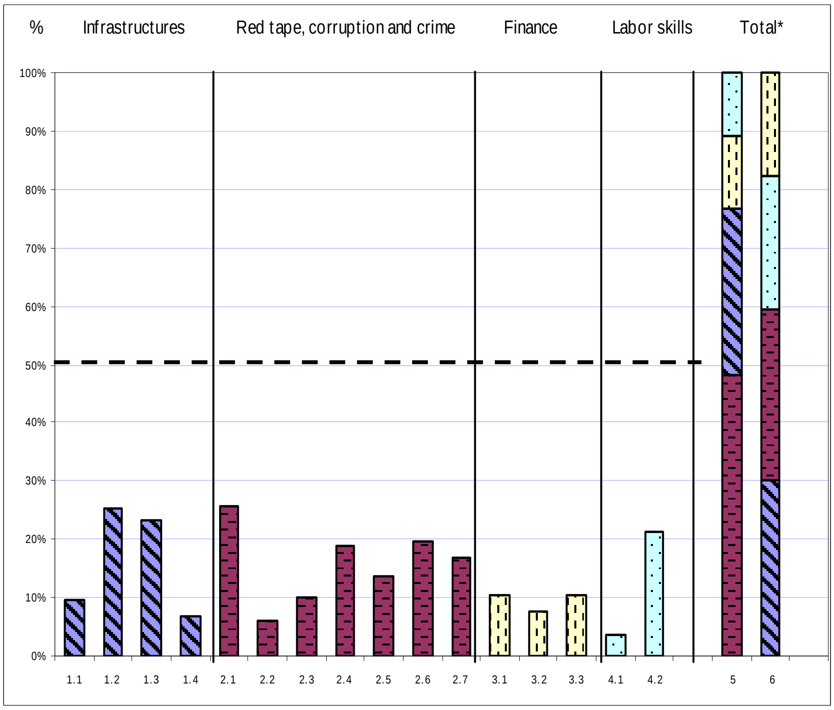
Note: Contributions computed according to section 5.8. The productivity measure used is the restricted Solow residual.
Source: Author's calculations with India ICS data.

Figure 5.10: Weight of each block of IC variables on the sample means of economic performance measures. (Software and ITES sub-sector)



Note: The weight of each block or group of IC variables to the sample means comes from the contributions of Figures5.6 to 5.9. We take the percentage contributions in absolute value and we compute the relative weight of each block. For the case of productivity we take the IC contributions to average log-productivity from Figure 4.6.
Source: Authors' calculations with India ICS data.

Figure 6: Firm's perceptions; percentage of firms that considers each one of the following problems as a severe obstacle to firms' economic performance.



1. Infrastructures.
1.1 Telecommunications
1.2 Electricity
1.3 Transportation
1.4 Customs and trade regulations

2 . Red tape, informality and others.
2.1 Corruption
2.2 Crime, theft and disorder
2.3 Anti-competitive or informal practices
2.4 Regulatory policy
2.5 Tax rates
2.6 Tax administration
2.7 Business licensing and operating permits

3. Finance.
3.1 Access to land
3.2 Cost of finance
3.3 Macroeconomic uncertainty

4. Labor skills.
4.1 Labor regulations
4.2 Skills and education of available workforce

5. Total relative weights.
6. Average group relative weights.
* (Totals are computed as the relative weigh of each group of perceptions over the sum of all perceptions' weights)

Source: Authors' calculations with India ICS data.

Chapter IV

Economic Determinants of India`s Investment Climate

Retail Sector

Economic Determinants of India`s Investment Climate

Retail Sector

Abstract

India has one of the highest retail densities in the world. The retail sector in India is composed of small retail stores at one extreme and large modern format stores at the other. This sector is characterized by low labor-productivity and modest employment growth. In recent years, local and national administrations have adopted a number of policies to facilitate the growth of the sector, particularly of the modern format stores. In this paper, we develop a specific econometric methodology to identify key investment climate constrains that slow down growth of the sector, concretely for the case of modern format stores. Since retail industry is the second largest provider of jobs in the economy of India, also we evaluate the IC determinants of demand for labor in this sector. Our main findings shows that the principal obstacles what Indian retail stores face are unreliable infrastructures, access to finance and bureaucracy.

JEL Clasification: D24; D61; J20; J30; L81; O40; O53

Keywords: *labor-productivity, demand for labor, investment climate, average contribution, modern format stores.*

Table of Contents

1. INTRODUCTION....177

2. ANALYSIS OF LABOR-PRODUCTIVITY....181
2.1 Traditional Fast Moving Consumer Goods (FMCG) Stores ____ 185
2.2 Modern Format Stores____ 185
2.3 Consumer Durable Stores ____ 186

3. ANALYSIS OF EMPLOYMENT DEMAND 186
3.1 Employment Demand for Micro and Small Stores ____ 188

4. CONCLUSIONS....189

REFERENCES....191

APPENDIX A: ECONOMETRIC MODELS 194
A.1 Labor Demand Equation: Conditional Labor Demand ____ 194
A.2 Labor Productivity Equation ____ 196
A.3 Economic Interpretation of Individual Coefficients of the Employment Demand Conditional on the Level of Sales.____ 196
A.4 Economic Interpretation of Individual Coefficients of the Employment Demand Conditional on Labor-Productivity. ____ 199

APPENDIX B: DATA TRANSFORMATIONS 202

APPENDIX C: DEFINITIONS....205

APPENDIX D: TABLES AND FIGURES....208

1. Introduction

Developing countries are increasingly concerned about improving country competitiveness and productivity, as they face the increasing pressures of globalization and attempt to improve economic growth and reduce poverty. Among such countries, Investment Climate (IC) have become a standard instrument for identifying key obstacles to country competitiveness and imputing their impact on productivity, in order to prioritize policy reforms for enhancing competitiveness.

A significant component of country competitiveness is having a good investment climate or business environment. The investment climate, as defined in the WDR (2005), is “the set of location-specific factors shaping the opportunities and incentives for firms to invest productively, create jobs and expand.” It is now well accepted and documented, conceptually and empirically, that the scope and nature of regulations on economic activity and factor markets - the so-called investment climate and business environment - can significantly and adversely impact productivity, growth and economic activity (see Bosworth and Collins, 2003; Escribano and Guasch (2005 and 2008a), Escribano et al. (2008b and 2008c), Rodrik and Subramanian, 2004; McMillan, 1998 and 2004; OECD, 2001; Djankov et al., 2002; Haltiwanger, 2002; He et al., 2003; World Bank, 2003; and World Bank, 2004 a,b).

Government policies and behavior exert a strong influence on the investment climate through their impact on costs, risks and barriers to competition. Key factors affecting the investment climate through their impact on costs are: corruption, taxes, the regulatory burden and extent of red tape in general, factor markets (labor, intermediate materials and capital), the quality of infrastructure, technological and innovation support, and the availability and cost of finance.

For example, Kasper (2002) shows that poorly understood “state paternalism” has usually created unjustified barriers to entrepreneurial activity, resulting in poor growth and a stifling environment. Kerr (2002) shows that regulation is a massive deterrent for investment and economic growth. As a case in point, McMillan (1998) argues that obtrusive government regulation before 1984 was the key issue in New Zealand’s slide in the world per-capita income rankings. de Soto (2002) describes one key adverse effect of significant business regulation and weak property rights: with costly firm regulations, fewer firms choose to register and more become informal. Also, if there are high transaction costs involved in registering property, assets are less likely to be officially recorded, and therefore cannot be used as collateral to obtain loans, thereby becoming “dead” capital.

Likewise, poor infrastructure and limited transport and trade services increase logistics costs, rendering otherwise competitive products uncompetitive, as well as limiting rural production and people's access to markets, which adversely affects poverty and economic activity (Guasch 2004).

The report of Mckinsey (2001) predicts that reforms in India's retail sector could create 8 million jobs until 2010 if the efficiency gains, of modern format stores, are transmitted to consumers through lower prices, and create higher productivity and restructuring of the upstream supply chain (like in labor-intensive food processing industries). For that they recommend; foreign direct investment (FDI) should be allowed in the retail sector, artificial scarcity land market barriers should be removed, bureaucracy should be reduced and operational disadvantages faced by supermarkets, versus counter stores, should be removed by having more flexible labor laws and tax policies.

The Economist (2006) mentioned that most Indian shops belong to the "unorganized retailing" (small, family owned shops surviving on unpaid labor and often free land for small stall) but "organized retailing", even if it only accounts for 2-3% of the total, is growing at 20% per year and shopping malls are springing up in very big city. There has been some progress with respect to FDI. In January 2006 the government relaxed some of the rules on FDI so that foreign firms would be allowed to hold up 51% in single-brand retail outlets and in multiproduct single-brands (supermarket own brands would be allowed) stores would also be allowed to invest. In India private consumption accounts for 64% of gross domestic product (GDP) (more than Europe which is 58%), GDP is growing fast (7.5% during the last three years) and has one of the youngest populations, therefore the prospects for the retail sector growth are very optimistic if the government is able to relax the constraints on FDI.

In the case of developing countries, inward FDI may increase productivity, simply because foreign investors, often based in more advanced economies, dispose of more productive technologies. In this case, domestically owned and foreign owned firms get their productivity from different exogenous distributions. To use the words of Evenett and Voicu (2002), are foreign investors picking winners or creating them? The positive contemporaneous correlation between foreign ownership and productivity also holds up when one focuses on FDI between developed countries. The recent theoretical work of Helpman, Melitz and Yeaple (2004) proposes a mechanism, similar to the one in Melitz (2003) that rationalizes this fact. Because of the fixed costs involved in setting up an affiliate plant abroad, only the most productive firm are able to become multinationals.

What keeps low productivity firms from exporting in both Melitz (2003) and Yeaple (2005) is the existence of a fixed cost to enter export markets. There is empirical evidence supporting this view.

The conventional wisdom associates foreign direct investment with higher productivity. According to Markusen (1995), one important stylized fact is that multinationals are prevalent in firms and industries with high levels of R&D, a large share of professional and technical workers, and products that are new and/or technically complex. This is in line with Dunning (1993) who argues that to overcome local barriers, multinationals must have some intangible assets, such as superior technologies or more advanced management techniques. Markusen (1995) refers to this as knowledge-based assets.

The pursuit of greater competitiveness and a better investment climate is leading countries - often assisted by multilaterals such as the World Bank - to undertake their own studies to identify the principal bottlenecks in terms of competitiveness and the investment climate, and evaluate the impact these have, to set priorities for intervention and reform. The most common instrument used has been firm-level surveys, known as Investment Climate (ICs), from which both subjective evaluations of obstacles and objective hard-data numbers with direct links to costs and productivity are elicited and imputed. Such surveys collect data at firm level on the following themes: a) infrastructure, b) red tape, informality and others, c) finance and corporate governance, d) quality, innovation and labor skills and d) other control variables like legal status of the firm, age and size of the firm, etc.

While the Investment Climate Assessments are quite useful in identifying major issues and bottlenecks as perceived by firms, the data collected is also meant to provide the basic information for an econometric assessment of the impact or contribution of the investment climate (IC) variables on productivity. In turn, that quantified impact is used in the advocacy for, and design of, investment-climate reform. Yet providing reliable and robust estimates of productivity estimates of the IC variables from the surveys is not a straightforward task since; first, the surveys do not provide panel-type data on IC variables; second, neither the production function parameters nor the functional form are observed; and third, there is an identification issue separating total factor productivity (TFP) component from the inputs of the production function.

When any of the production function inputs is influenced by common causes affecting productivity, like IC variables or other plant characteristics, there is a simultaneous equation problem. In general, one should expect the productivity to be correlated with the production function inputs (technological progress is not Hicks neutral) and, therefore, inputs should be treated as endogenous regressors when estimating production functions. This property has demanded special care with the econometric specification when estimating those productivity effects and in the choice of the most appropriate way of measuring productivity. There is an extensive literature discussing the advantages and disadvantages of using different statistical estimation techniques and/or growth accounting (index number) techniques to estimate productivity. For overviews of different productivity concepts and aggregation alternatives based on firm level data see, for example, Foster, Haltiwanger and Krizan (1998), Batelsman and Doms (2000) and Olley and Pakes (1996).

We believe that improving the investment climate (IC) is a key policy instrument to promote economic growth and to mitigate the institutional, legal, economic and social factors that are constraining the fats convergence of per capita income and labor productivity to more developed countries. For that, we need to identify the main investment climate variables that affect economic performance measures like labor-productivity and employment and this is the main goal of this paper.

The econometric analysis of the investment climate (IC) effects on productivity and employment of the retail sector of India, is based on the World Bank investment climate survey (ICs) done by in 2006 covering firm level information of fiscal year 2005-06. Investment climate (IC) variables are grouped in five blocks: *i) Infrastructure* (with 5 variables), *ii) Red tape, informality and others* (11 variables), *iii) Finance and Accountability* (13 variables), *iv) Quality, Innovation and Labor Skills* (2 variables) and *v) Other Contro*l variables (8 variables).

From the analysis of *firm perceptions* about the *most serious obstacles* of the retail sector (see Figures 6.1 and 6.2 of the appendix D), we identify the following patterns: 30% of the firms mention that the most serious obstacles are related to IC variables from three groups: i) infrastructure, ii) red tape, informality and others and iii) finance and accountability. Labor skills are the most serious obstacles for only 10% percent of the firms.

A more disaggregated analysis of firm's perceptions based on the most serious obstacles for firm growth (see Figure 6.2 of the appendix), allow us to obtain the following conclusions: within the block of infrastructures IC variables the main component is *electricity* (29%). Within red, tape, corruption and crime, *corruption* (11%) is an important factor, followed by *high tax rates* (9%) and *crime* (4%). Within the group of finance and accountability access to finance is the most serious IC factor (16%).

With the econometric analysis of sections 2 and 3 we are able to check if the empirical results are consistent with firm's perceptions on IC bottlenecks for growth. For that purpose, we performed a detailed empirical analysis of labor-productivity and employment at the firm level based on the IC survey of the retail sector of India. The maximum number of firms available in the investment climate survey of the retail sector of India is 1948, but due to the existence of missing values, zeros and outliers, we ended up having around 1897 (see Tables B of the appendix D and the section on data transformations of appendix B). The problem of missing values is very serious as can be seen from Tables B of appendix D and the implications in terms of missing IC variables are clear from

Table A. Notice that only those IC variables of Table A of appendix D that are in bold letters are available for the econometric analysis.

The *productivity analysis* is done for the whole retail sector as well as for the three main types of stores; traditional FMCG, modern format stores and consumer durable stores. The empirical results are presented in section 2.

The *employment analysis* is done at the aggregate retail sector level and for each category of firm size. However, for estimation purposes, in order to have enough observations in each group, the initial five size categories were grouped in two groups of stores: *whole retail sector* and *micro and small* (representing 85% of the stores). Most of the micro and small firms are from the traditional FMCG and the consumer durables stores, while most of the medium, large and extra-large firms are from the modern format stores. The empirical results from the econometric analysis on labor demand are presented in section 3.

2. Analysis of Labor-Productivity

The productivity analysis is based on the concept of labor-productivity, instead of total factor productivity[26] (TFP), due to a lack of information on some of the basic inputs of the production functions, like the capital stock. Therefore, the analysis of the retail sector of India is based on reduced form of the production function. In particular, the log of labor productivity is explain by the log of wages, the log of rental cost of capital, the investment climate (IC) variables and other control variables, including dummy variables for industry (traditional FMCG, modern format stores and consumer durable stores) for regions (north, south, east and west) and for firm size (micro, small, medium, large and extra-large).

The *reduced form* equation of labor productivity in terms of the input prices (w, r), investment climate (IC) variables and other control (C) variables, see equation (A.11) of appendix A, is given by;

$$\log\left(\frac{Y}{L}\right)_{it} = \gamma_0 + \gamma_1 \log w_{it} + \gamma_2 \log r_{it} + \gamma_3' IC_i + \gamma_4' C_{it} + \varepsilon_{i,YL} \qquad (2.1)$$

[26] Notice that in most of the ICs done on the manufacturing sector, the analysis was based on TFP and not on labor productivity. See for example Escribano and Guasch (2005).

where i=1,…, 1897 an t=2006. This is the *labor productivity equation* that we estimate in Table C of Appendix D. Notice that we are treating the investment climate (IC) variables as *observable fixed effects*, as was suggested in Escribano and Guasch (2005).

Under general conditions, the ordinary least squares (OLS) estimator is a consistent estimator of the parameters of equation (2.1), after controlling for observable fixed effects (IC_i). To control for the usual *heteroskedasticity* of firm level data, we use robust standard errors. The main remaining problem of OLS is the possible *endogeneity* of certain IC variables. In those cases we have evidence of endogenous IC variables we substitute the endogenous IC_i variable for their corresponding region-industry average ($\overline{IC}$). This procedure has the additional advantage of recovering for the econometric analysis many IC variables with missing values; see Escribano and Guasch (2005 and 2008a) for more details on this econometric methodology applied other countries based on IC surveys.

In order to evaluate the contribution of each explanatory variable to labor-productivity we evaluate equation (1) at the sample mean. This decomposition is exact when the sample mean of the residuals is zero, $\bar{\hat{\varepsilon}}_{itYL} = 0$ since we have a constant term in regression (2.1)

$$100 = \hat{\gamma}_0 \left(\frac{1}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_1 \left(\frac{\overline{\log}\, \overline{w}}{\overline{l\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_2 \left(\frac{\overline{\log}\, \overline{r}}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_3' \left(\frac{\overline{IC}}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 + \hat{\gamma}_4' \left(\frac{\bar{C}}{\overline{\log}\left(\frac{Y}{L}\right)} \right) 100 \quad (2.2)$$

The empirical results on *labor-productivity* are the following: Explaining labor-productivity the main single variable is wages (see Table C.I of appendix D). The wage elasticity is 0.34 which says that a 1% increase in wages, keeping the rest of the variables constant, will produce an increase labor productivity of 0.34%. However, the impact of average wages (in logs) on average labor-productivity is very high, equal to 85.8%. The reason is the following; when there is an increase in wages there will be an increase in capital deepening (capital labor ratio) and an increase in total factor productivity (TFP), since the employees have more incentives (make more effort) to increase their individual productivity, and therefore the total impact on average labor productivity of the firm can be very high.

The estimated IC equation of the retail sector is able to explain 35% of the variability of labor productivity (R^2=0.35). For the marginal effects of IC variables see Figure 4.1. In terms of investment climate variables, the most important block of IC variables is finance and accountability which represent 30.5% of the total IC impact on *average labor productivity* (in logs). The second block is quality, innovation and labor skills with 19.5%, followed by infrastructures with 18.1%, other control variables with 16.7% and then red tape, informality and others with 15.2% each. Those investment climate patterns are preserved at the aggregate retail sector level with very similar proportions over *aggregate labor productivity* (in logs). (See Figure 4.5)

From the finance and accountability block two IC variables are the most important ones; having a *current account* and having an *external auditory* of the financial annual statements (accountability) of the stores, see Figure 4.1. Those results are consistent with firm perceptions that identify *access to finance* as one on the main obstacles for firm growth, see Figure 6.1 of appendix D.

The *experience of the manager and dummy for computer* are the only two IC variable of the block of quality, innovation and labor skills and it is very important since it represents 19.5% of the total IC impact on average labor productivity, see Figures 4.5 of appendix D. However, these variables were not identified as important constraint by firm's perceptions.

Within the block of infrastructures the *duration of the power outages* and the number of days of *inventories* the fact are the main IC factors; see Figure 6.2. This is consistent with firm's perceptions that identify electricity as the main obstacle for firm growth.

Regarding other control variables group, only one variable has a significant contribution on average labor productivity, the age of the firm. Its individual impact is one of the largest among IC variables.

Within the block of red tape, corruptions and crime, the main three IC variables are; the *store visited by any agency*, security and *criminal attempt;* see Figure 4.1. This is consistent with the firm perceptions that identify *corruption and crime, theft and disorder* as one on the main obstacles for firm growth; see Figure 6.1.

The Olley and Pakes (1996) decomposition of productivity, (O&P), has two elements; the average productivity and the efficiency or covariance term. As will become clear later on, the O&P decompositions would allows us to do a sector by sector, region by region, etc., evaluation of the impact of IC variables on average productivity and on efficiency without requiring further parameter estimation with fewer observation.

Let $\left(\frac{Y}{L}\right)_{jt} = \sum_{i=1}^{N_{jt}} s_{j,it}^{Y} \left(\frac{Y}{L}\right)_{j,it}$ be the aggregate labor-productivity of industry j at time t obtained as the weighted average of i-plant-level productivity $\left(\frac{Y}{L}\right)_{j,it}$ in sector j at year t, where N_{jt} is the number of firms in sector j where j = 1, ... ,8 at time t. The weights ($s_{j,it}^{Y}$) indicate the share of sales of firm i in year t over the total sales (Y) of sector j of that year ($s^Y_{j,it} = \frac{Y_{j,it}}{\sum_{i=1}^{N_{j,t}} Y_{j,it}}$). Let $\left(\frac{Y}{L}\right)_{jt} = \frac{1}{T}\sum_{i=1}^{N_{jt}}\left(\frac{Y}{L}\right)_{j,it}$ be the sample average labor-productivity of the firms of sector j in year t. Then the annual aggregate labor-productivity of industry j can be decomposed as in (3) where $\tilde{s}^Y_{j,it} = (s^Y_{j,it} - \bar{s}^Y_{j,t})$ and $\left(\frac{\tilde{Y}}{L}\right)_{j,it} = \left(\frac{Y}{L}\right)_{j,it} - \left(\frac{\bar{Y}}{L}\right)_{j,t}$ are in deviations to the mean. Notice that $\bar{s}^Y_{j,t} = \frac{1}{N_{jt}}$ the Olley and Pakes (1996) decomposition (O&P) of labor-productivity in levels is:

$$\left(\frac{Y}{L}\right)_{jt} = \left(\frac{\bar{Y}}{L}\right)_{jt} + \sum_{i=1}^{N_{jt}} \tilde{s}^Y_{j,it}\left(\frac{\tilde{Y}}{L}\right)_{j,it} . \tag{2.3}$$

The first term, $\left(\frac{\bar{Y}}{L}\right)_{jt}$, is the *average labor-productivity* of industry j in year t and the second term ($\sum_{i=1}^{N_{jt}} \tilde{s}^Y_{j,it}\left(\frac{\tilde{Y}}{L}\right)_{j,it}$) = $N_{jt}\,\mathrm{cov}(s^Y_{j,it}, \left(\frac{Y}{L}\right)_{j,it})$, measures the *allocative efficiency* or covariance between the share of sales and labor-productivity, $\mathrm{cov}(s^Y_{j,it}, \left(\frac{Y}{L}\right)_{j,it})$, multiplied by the number of firms, N_{jt}, that belong to sector j in year t. If the efficiency term is positive, then the larger it is, the higher will be the share of sales that goes to more productive firms, allocation efficiency is increased and sector j productivity is enhanced. If the efficiency term is negative, there are allocation inefficiencies since the more negative the efficiency term is, the higher will be the share of sales that goes to less productive firms, reducing sector j labor-productivity.

We can compute the aggregate labor-productivity of sector j but *in mixed decomposition*. Let $\log\left(\frac{Y}{L}\right)_{jt} = \sum_{i=1}^{N_{jt}} s^{\log Y}_{j,it} \log_{j,it}\left(\frac{Y}{L}\right)$ be the aggregate log labor-productivity of industry j at time t obtained as the weighted average of i-plant log labor-productivity in sector j at year t, where N_{jt} is the number of firms in sector j where j = 1, ... ,8. The weights ($s^{\log Y}_{j,it}$) indicate the share of firm i sales (Y) in year t over the total sales of sector j of that year. We therefore have equivalent labor-productivity decomposition, equation (2.3), but with variables in mixed decomposition.

Aggregate labor-productivity of the retail sector can improve because there is an increase in average labor productivity or an improvement in allocative efficiency; see Figure 2 of appendix D. The allocative efficiency component of the retail sector of India is positive and tells us that on average the firms that win market shares become also more productive. In terms of this *allocative* *efficiency component* the main IC block is Finance and Accountability with an even higher IC impact representing 40% of the total efficiency, followed by quality, innovation and labor skills with 29% and red tape, informality and others with 20.3% while the other IC blocks (infrastructures and other control variables) contribute 7% and 3% respectively, see Figure 4.5 of appendix D.

From the Olley and Pakes (1996) decomposition of aggregate labor productivity into average labor-productivity and allocative efficiency, we obtain the following result (see Figure 2 and Figure 3 of appendix D); the most productive firms are the medium size stores, followed by the small and micro. The least productive stores are the extra-large that at the same time are, together with large stores, the ones that suffer more the effects of the investment climate. This is especially relevant in terms of the IC blocks of Finance and Accountability, infrastructures and red tape, informality and others. However by regions, north, south, east and west, there is no significant difference. By type of stores, the most productive ones are the consumer durable stores and the modern format stores followed by the traditional FMCG, see Figure 2 and Figure 3 of appendix D.

In summary, the econometric analysis of the labor productivity of the retail sector is able to identify certain priorities for policy recommendations in the retail sector of India. In particular, they are related to reducing the financial constraints, improving the access to finance and at the same time improving the abilities and the skills (training) of the managers in order to reduce the time to get the required experience of the mangers. Other important IC aspects for designing specific policies are related to reductions in corruption and crime and improvements in the quality of the electricity system.

2.1 Traditional Fast Moving Consumer Goods (FMCG) Stores

The estimated IC equation of the traditional fast moving consumer goods (FMCG) sector is able to explain 23% of the variability of labor productivity (R^2=0.23) which is lower than the 35% of the whole retail sector; see the second column of Table C.I of appendix D.

The impacts of IC variables on labor productivity by blocks are similar to the impacts on the whole retail sector; see Figure 4.5.

2.2 Modern Format Stores

The estimated IC equation of the modern format sector is able to explain 44% of the variability of labor productivity (R^2=0.44) which is higher than the 30% of the whole retail sector; se Table C of appendix C.

In this type of stores, the impacts of IC variables on labor-productivity by blocks are different than the impacts on the whole retail sector; see Figure 4.5. The *main increases are in the labor* *skills and in the infrastructure blocks*. The experience of the manager increases its labor productivity impact from 13.5% of the whole retail sector to 50% of the modern format stores. The main block of IC variables that losses relevance is the finance and accountability that is reduced from 33% to 1.2%. The results are very also very similar for aggregate productivity.

In terms of *allocative efficiency*, the main increases are in the IC groups of red tape, informality and others and in labor skills; see Figure 4.5. The increase in red tape, informality and others went from 20% to 39%. Once again, the block that reduced its relative impact is finance and accountability.

2.3 Consumer Durable Stores

The estimated IC equation of consumer durable stores is able to explain 31% of their corresponding labor productivity (R^2=0.31), very similar to the whole retail sector; see Table C.I of appendix D. The main increase of importance is in the other control variables group. (See Figure 4.5)

In terms of *allocative efficiency,* the main differences are in terms of the increases in the IC block of finance and accountability going from 40.6% of the whole retail sector to 54.6% of the consumer durable stores; see Figure 4.5. The *increase in the relative importance of the access to finance* variables is at the cost of eliminating the effect of the increases in the number of days of inventories in this type of stores.

3. Analysis of Employment Demand

The Economist (2006) mentioned that India's retail industry is the second largest provider of jobs (the first is agriculture) representing 6-7% of total employment. Therefore, employment demand in the retail sector is a very sensitive political issue.

The *conditional employment demand* is usually express in terms of the level of output that the firm wants to produce, total factor productivity (TFP), input prices (wages, the rental cost of capital, etc.) and investment climate (IC) variables and other control (C) variables. Since TFP cannot be estimated in the retail sector of India, because of the lack of observations on the capital stock, we obtain a *semi-reduced form* of the conditional employment demand by substituting TFP by its determinants in terms of IC and C variables.

$$\log L_{it} = \beta_0 + \beta_1 \log\left(\frac{Y}{L}\right)_{it} - \beta_2 \log w_{it} + \beta_3 \log r_{it} + \beta_4' IC_i + \beta_5' C_{it} + \varepsilon_{L,it}\,. \qquad (3.1)$$

This is the equation we estimate in Table D of appendix D. The estimation procedure is two stage least squares (2SLS) since labor productivity (Y/L) and wages are correlated with the error term (ε_L) of (3.1). Remember that we are considering a firm that produces good Y with two inputs, labor (L) and Capital (K) with a Cobb-Douglas technology with a total factor productivity component (TFP),

$$Y_{it} = TFP_{it} \, K_{it}^{\ \alpha} \, L_{it}^{\ \beta} \, . \tag{3.2}$$

Both $\alpha \; and \, \beta$ are positive and if $\alpha + \beta = 1$ there are *constant returns to scale* (CRS), if $\alpha + \beta$ <1 there are *decreasing return to scale* (DRS) and if $\alpha + \beta$ >1 *increasing return scale*. To interpret the influence of labor productivity (Y/L) and wages (w) on employment (L), it is important to remember that the sign of the coefficients are related with the type of return to scale $\beta_1 = \left(\frac{1}{\alpha + \beta - 1} \right)$ and $\beta_2 = \left(\frac{\alpha}{\alpha + \beta - 1} \right)$, see equation (A.7) of appendix A.

In order to evaluate the contribution of each explanatory variable to the labor demand we evaluate equation (3.1) at the sample mean. This decomposition is exact when the sample mean of the residuals is zero, $\bar{\hat{\varepsilon}}_{L,it} = 0$ since we have a constant term in regression (3.1).

$$100 = \hat{\beta}_0 \left(\frac{1}{\overline{\log L}} \right) 100 + \hat{\beta}_1 \left(\frac{\overline{\log \left(\frac{Y}{L} \right)}}{\overline{\log L}} \right) 100 + \hat{\beta}_2 \left(\frac{\overline{\log \bar{w}}}{\overline{\log L}} \right) 100 + \hat{\beta}_3 \left(\frac{\overline{\log \bar{r}}}{\overline{\log L}} \right) 100 + \hat{\beta}_4 \left(\frac{\overline{IC}}{\overline{\log L}} \right) 100 + \hat{\beta}_5 \left(\frac{\bar{C}}{\overline{\log L}} \right) 100 \tag{3.3}$$

Our conditional labor demand will be expressed in terms of *labor productivity* and not in terms of the level of output of the firm. This has the advantage that the signs of the coefficients of labor productivity and wages depend on the type of *returns to scale* of the firm, see appendix A, and will become clear in the empirical analysis.

From the results of estimating the conditional labor demand for the whole retail sector of India we get that *there are increasing returns to scale* since the coefficient of labor productivity is positive, the coefficient of wages is negative and the one of the rental cost of capital is positive; see the first column of Table D of appendix D.

In particular, the elasticity of employment with respect to labor productivity is 0.25, meaning that a 1% increase in labor productivity, keeping the rest of the variables constant, will generate a 0.25% increase in the employment hours. The impact of average labor productivity (in logs) on average employment (in logs) is 37%.

The elasticity of employment with respect to wages is -0.47, meaning that a 1% increase in wages, keeping the rest of the variables constant, will reduce the amount of employment by 0.47%. However, the impact of average wage (in logs) on average number of hours (in logs) is equal to -617%.

The elasticity of the rental cost of capital on employment is 0.15, meaning that a 1% increase in the rental cost of capital, keeping the rest of the variables constant, will generate a 0.15% increase in the employment hours. The impact of the average rental cost of capital (in logs) on average employment (in logs) is 198%.

In terms of investment climate variables, the most important block of IC variables is other control variables which represent 50% of the total IC impact on *average labor demand* (in logs); see Figure 5.3 of appendix D. The second block is infrastructures with a 20%, followed by finance and accountability with 14% and by quality, innovation and labor skills with 7.4%.

Within the main block is other control variables, see Figure 5.1, the only relevant variable is the age of the firm. The following block is infrastructures with single variable, days *of inventory*, representing 22% of average employment (in logs). Within the block of finance and accountability, the two variables with significant impact are: dummy for overdraft and dummy for external auditory.

As mentioned by The Economist (2006), in India some sorts of food processing are “reserved” for small business and in towns the size of shops and permissible opening hours are governed by local rules. Those size-dependent policy restrictions can have very large negative impact even at the macroeconomic level, see Guner, Ventura and Xu(2006) and Mohan(2002). We will therefore evaluate the employment differences of IC variables by size of the stores.

3.1 Employment Demand for Micro and Small Stores

The group of micro and small stores *has increasing returns to scale* since the coefficient of labor productivity is positive, the coefficient of wages is negative and the one of the rental cost of capital is positive.

In particular, the elasticity of employment with respect to labor productivity is 0.21, meaning that a 1% increase in labor productivity, keeping the rest of the variables constant, will generate a 0.21% increase in the employment hours; see the second column of Table D of appendix D. The impact of average labor productivity (in logs) on average employment (in logs) is 34%.

The elasticity of employment with respect to wages is -0.465, meaning that a 1% increase in wages, keeping the rest of the variables constant, will reduce the amount of employment by

0.465%. However, the impact of average wage (in logs) on average number of hours (in logs) is equal to -686%, much higher that for the whole retail sector.

The elasticity of the rental cost of capital on employment is 0.08, meaning that a 1% increase in the rental cost of capital, keeping the rest of the variables constant, will generate a 0.08% increase in the employment hours. The impact of the average rental cost of capital (in logs) on average employment (in logs) is 125%.

In terms of investment climate variables, the relative impact of each block of IC variables changes dramatically with respect to the whole retail sector; see Figure 5.3. Now, the most important block of IC variables is red tape corruption and crime due to the impact of *overdue payments* which is 59%; see the central columns of Table D. *Finance and accountability* represents now 30% of the total IC impact on *average labor demand* (in logs). Four important IC variables appear within this block of finance; having a credit line, having a current account, using internal funds as financing sources and having an external auditory of the financial statements. The last block is quality, innovation and labor skills with a 1.5%. For the individual IC effects, Figure 5.5 shows that the *micro stores* are the ones that suffer more the effects of the investment climate on employment and specially those IC variables of the group of *finance*.

4.Conclusions

This paper develops specific econometric methodology to identify key investment climate factors that constrain labor-productivity on retail sector in India. Additionally, we evaluate the IC determinants of demand for labor in this sector. The retail sector in India is composed of small retail stores at one extreme and large modern format stores at the other. Sector what is largely fragmented and is dominated by traditional retail sector.

Although the retail sector has been growing, is a sector characterized by low labor-productivity and modest employment growth. In recent years, local and national administrations have adopted a number of policies to facilitate the growth of the sector, particularly of the modern format stores.

Our main econometric results are exposed in Figures 4.1-4.5. Concretely, Figure 4.2 shows the relative contributions of each group of IC variables on mixed Olley and Pakes decomposition of aggregate labor-productivity for traditional FMCG stores. The results are compared with the results for whole sector (see Figure 4.5). Conclusions are very similar (traditional retail stores represented 97% of the total sector sales in 2005). The largest relative impacts come from finance and accountability group, with relative percentage contributions of 33.6 and 35.4 respectively. This structure of contributions is maintained in two components of Olley and Pakes decomposition, even finance and accountability variables have larger relative impact on the allocative efficiency.

For the case of modern format stores (see Figure 4.3), the structure of IC groups of variables relative contribution on mixed Olley and Pakes decomposition on aggregate labor-productivity is completely different than the case of whole retail sector. In this sub-sector, the largest contribution, clearly, is given by quality, innovation and labor skills variables, overall on the average labor-productivity component (50%). On the other component of Olley and Pakes decomposition (allocative efficiency) the largest impact comes from red tape, informality and others with more than 39% of total contribution.

Finally Figure 4.4 summarizes the results of the same analysis for consumer durable stores. The main finding is the lost of importance of quality, innovation and labor skills factors on aggregate productivity for this sub-sector.

Figure 4.1 illustrates the individual contribution of each significant IC variables on mixed Olley and Pakes decomposition on aggregate labor-productivity for whole retail sector. Dummy for current account, the age of the firm and the experience of the manager are the IC variables with largest individual impact on average productivity. The highest contributions on allocative efficiency are given by dummy for computer, dummy for current account and dummy for external audit.

Figure 5.3 compares the relative importance of groups of IC and C variables in terms of contributions to average log-employment between the whole retail sector and micro&small firms.

For the case of whole retail sector, other control variables is the group with largest contribution following by infrastructures. However, in the analysis of micro&small firms, the two groups with most relevant impacts are: red tape, informality and others and finance and accountability, respectively.

Figure 5.1, for whole retail sector, and Figure 5.2, for micro&small firms, show the individual contributions on average log-employment. Similar conclusions we can obtain in the case on *real wages* impact on average log-employment. Real wages has the largest relative impact of the whole contribution of IC and C variables to average log-employment.

References

Amin Mohammad (2006). “Labor Regulations and Technology in India´s Retail Stores”. Mimeo, The World Bank, Washington DC.

Amin Mohammad (2008). “Retailing in India: Assesing the Investment Climate”. Mimeo, The World Bank, Washington DC.

Bartelsman E. J. and M. Doms (2000). “Understanding Productivity: Lessons from Longitudinal Microdata”. *Journal of Economic Literature*, Vol. 38, Nº 3, 569-594.

de Soto, Hernando (2002). “The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else”. New York: Basic Books Press.

Djankov, S. and Hoekman, B., (2000). Foreign Investment and Productivity Growth in Czech Enterprises, *World Bank Economic Review*, 14, 49-64

Djankov, Simeon, La Porta, Rafael, Lopez-de-Silanis, Florencio, and Shleifer, Andrei (2002). “The Regulation of Entry”. *Quarterly Journal of Economics* 117, February 2002. 1-37.

Dobson Consulting (1999). “Buyer Power and Its Impact on Competition in the Food Retail Distribution Sector of the European Union”. Report prepared for the European Commossion-DGIV, 1-197.

Dunning. J.H., (1993). Multinational Enterprises and the Global Economy, Addison Wesley.

Escribano Alvaro and J. Luis Guasch (2005); “Assessing the Impact of the Investment Climate on Productivity using Firm Level Data: Methodology and the Cases of Guatemala, Honduras and Nicaragua,” *World Bank Policy Research Working Paper* # 3621, The World Bank, June.

Escribano A. and J. L. Guasch (2008); “Robust Methodology for Investment Climate Assessment on Productivity: Application to Investment Climate Surveys from Central America,” *Working Paper # 08-19 (11)*, *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena. (2008a); “Investment Climate and Firm’s Performance: Econometric and Applications to Turkey’s Investment Climate Survey,” Working Paper # 08-20 (12), *Universidad Carlos III de Madrid*.

Escribano, A., J. L. Guasch, M. de Orte and J. Pena (2008b); “Investment Climate Assessment Based on Demeaned Olley and Pakes Decompositions: Methodology and Applications to Turkey’s Investment Climate Survey,” Working Paper # 08-20 (13), *Universidad Carlos III de Madrid*.

Escribano A., N. Peltier, L. Garrido and H. Sing (2005b). “The Impact of Infrastructure on Competitiveness in Latin America: A Firm Level Analysis Based on Investment Climate Assessments”. *Paper presented at the Joint Conference of the Interamerican Development Bank (IADB) and The World Bank (WB)*, June 6-7.

Evenett, S.J. and Voicu, A., (2001). Picking winners or creating them? Revisiting the benefits of FDI in the Czech Republic, *unpublished manuscript*.

Guner N., G. Ventura and Y. Xu (2006). “Macroeconomic Implications of Size-Dependent Policies”. *Working Paper*, Universidad Carlos III de Madrid.

Guruswamy M., K. Sharma, J. P. Mohanty and Th. J. Korah (2006). “FDI in India´s Retail Sector: More Bad than Good? Discussion Paper of the Centre for Policy Alternatives. New Delhi.

Foster L., J. Haltiwanger and C.J. Krizan (1998). “Aggregate Productivity Growth: Lessons from Microeconomic Evidence”. *NBER Working Paper* W6803.

Guasch, J. Luis. (2004). “An Assesment of Logistic Costs and of their Impact on Competitiveness”, World Bank, Washington, DC.

Haltiwanger, John (2002). “Understanding Economic Growth”. The Need for Micro Evidence”. *New Zealand Economic Papers* 36 (1), 33-58.

Harris, R. and Robinson, C., (2003). Foreign Ownership and Productivity in the United Kingdom: Estimates for U.K. Manufacturing Using the ARD, *Review of Industrial Organization*, 22, 207-223.

He, Kathy S., Morck, Randall and Yeung, Bernard (2003). “Corporate Stability and Economic Growth”. William Davidson Working Paper No. 553.

Helpman, E., Melitz, M.J. and Yeaple, S.R., (2004). Export versus FDI with heterogeneous firms, *American Economic Review*, 94, 300-316.

Kasper, Wolfgang, (2002), “Losing Sight of the Lodestar of Economic Freedom: A Report Card on New Zeland’s Economic Reform”, NZ Business Roundtable.

Kerr, Roger (2002). ”The Quagmire of Regulation” NZ Business Roundtable. http://www.nzbr.org.nz/documents/speeches/speeches200/the_quagmire_of_regulation.pdf.

Markusen, J., (1995). The boundaries of multinational firms and the theory of international trade, *Journal of Economic Perspectives*, 9, 169-189.

McKinsey&Company (2001). “India: The Growth Imperative”.
http://www.mckinsey.com/mgl/publications/india.asp

McMillan, John, (1998), “ Managing Economic Change: Lessons from New Zeland”, *The World Economy 21,* 827-43.

McMillan, John (2004), “A Flexible Economy? Entrepeneurship and Productivity in New Zeland”, Working Paper, Graduate Scholl of Business, Stanford University, Stanford, CA.

Melitz, M.J., (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, *Econometrica*, 71, 1695-1725.

Mohan R. (2002). “Small-Scale Industry in India: A Critical Evaluation”. In Kruger, A. (Ed.), Economic Reforms and Indian Economy. University of Chicago press, Chicago, Illinois.

Olley G. S.and A. Pakes (1996). “The Dynamics of Productivity in the Telecommunications Equipment Industry”. *Econometrica*, Vol. 64, 6, 1263-1297.

Organization for Economic Cooperation and Development (2001). *Businesses’ Views on Red Tape,* Paris, OECD. http://www1.oecd.org/publications/e-book/4201101E.PDF.

Rodrik, Dani, and Arvind Subramanian (2004). “From ‘Hindu Growth to Productivity Surge: The Mystery of the Indian Growth Transition”. Harvard University, Cambridge, Mass. Processed.

The Economist (2006). “Retailing in India: Special Report”. 69-71.

Wooldridge J. M. (2002). Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, Massachusetts.

World Bank (2005): “World Development Report 2005: A Better Investment Climate for Everyone,” *The World Bank and Oxford University Press*, Washington, D.C.

World Bank (2009): “Doing Business Report 2009: Comparing Regulation in 181 Economies,” *The World Bank Group*, Washington DC, USA.

World Bank 2007. “Unleashing India`s Innovation: Toward Sustainable and Inclusive Growth”. Washington, D.C. World Bank

World Bank 2009. “India`s Investment Climate: Voices of Indian Business”. Washington, D.C. World Bank

World Economic Forum, The (2008).”Global Competitiveness Report”

Yeaple, S.R., (2005). A Simple Model of Firm Heterogeneity, International Trade, and Wages, *Journal of International Economics*, 65, 1-20.

Appendix A: Econometric Models

A.1 Labor Demand Equation: Conditional Labor Demand

Consider a firm that produces good Y with two inputs, labor (L) and Capital (K) with a Cobb-Douglas technology with total factor productivity (TFP) component. That is,

$$Y = TFP\,K^{\alpha} L^{\beta}\,. \tag{A.1}$$

Both $\alpha\, and\, \beta$ are positive and if $\alpha + \beta = 1$ there are *constant returns to scale* (CRS), if $\alpha + \beta \prec 1$ there are *decreasing return to scale* (DRS) and if $\alpha + \beta \succ 1$ *increasing return scale*.

The corresponding input prices are the *wages* (w) and the *rental cost of capital* (r) so that the *cost of producing* the Y units of output is (w L + r K).

When firms select the inputs corresponding to the *minimum cost* of production (Y), they set the optimal capital labor ratio (capital deepening) proportional to the relative price of the inputs,

$$\left(\frac{K}{L}\right) = \frac{\alpha}{\beta}\left(\frac{w}{r}\right) \tag{A.2}$$

The optimal demand for labor and capital should satisfy (A.1) and (A.2). In particular, solving the system for labor (L) we get the following *conditional labor demand of the firm*,

$$\log L = -\left(\frac{\alpha}{\alpha+\beta}\right)\log\left(\frac{\alpha}{\beta}\right) + \left(\frac{1}{\alpha+\beta}\right)\log Y - \left(\frac{1}{\alpha+\beta}\right)\log TFP - \left(\frac{\alpha}{\alpha+\beta}\right)\log w + \left(\frac{\alpha}{\alpha+\beta}\right)\log r\,. \tag{A.3}$$

Notice that since both $\alpha\, and\, \beta$ are positive, we get the expected signs of the coefficients (negative in w and positive in r).

A more interesting expression of the conditional labor demand is obtained in terms of labor productivity (Y/L) instead of the production level (Y) and is given by,

$$\log L = -\left(\frac{\alpha}{\alpha+\beta-1}\right)\log\left(\frac{\alpha}{\beta}\right) + \left(\frac{1}{\alpha+\beta-1}\right)\log\left(\frac{Y}{L}\right) - \left(\frac{1}{\alpha+\beta-1}\right)\log TFP - \left(\frac{\alpha}{\alpha+\beta-1}\right)\log w + \left(\frac{\alpha}{\alpha+\beta-1}\right)\log r$$

(A.4)

which can be simplified to

$$\log L = \beta_0 + \beta_1 \log\left(\frac{Y}{L}\right) - \beta_1 \log TFP - \beta_2 \log w + \beta_3 \log r\,. \tag{A.5}$$

It is clear that the coefficients of the variables log(Y/L) and log w have opposite signs and their corresponding sings depend of the nature of the returns to scale; β_1 and β_2 are positive under decreasing returns to scale (DRS) and β_1 and β_2 are negative under increasing returns to scale (IRS).

In the empirical application of the retail sector of India, we cannot estimate equation (A.5) since we do not have information on the capital stock (K) and we therefore cannot obtain TFP. However, under the usual condition that investment climate (IC) and other control (C) variables affect TFP,

$$\log TFP = \delta_0 + \delta_1' IC + \delta_2' C + \varepsilon_{TFP} \qquad \text{(A.6)}$$

substituting logTFP in (A.5), by their corresponding investment climate equation (A.6) , we get the following *semi-reduced form* of conditional labor demand,

$$\log L = \beta_0 + \beta_1 \log\left(\frac{Y}{L}\right) - \beta_2 \log w + \beta_3 \log r + \beta_4' IC + \beta_5' C + \varepsilon_L . \qquad \text{(A.7)}$$

Notice that in general the parameters β_4' should not only measure the indirect effect of IC on TFP since $\beta_4' \neq -\beta_1 \delta_1'$. We know from previous empirical evidence, see for example Escribano et al. (2006), that IC variables affect employment through TFP but also through a direct effect (say β_{IC}'). Therefore, we should expect the IC effects on employment demand (L) to be composed by two effects a direct (β_{IC}') effect of IC and indirect effect ($-\beta_1 \delta_1'$) of IC through TFP. Therefore, the coefficients of IC on employment demand are $\beta_4' = \beta_{IC}' - \beta_1 \delta_1'$.

We estimate the parameters (elasticities and semi-elasticities) of equation (A.7) in Table D of appendix D. The econometric estimation procedure is by *two stage least squares* (2SLS), since labor productivity and wages might be correlated with the error term (ε_L) of (A.7). For the economic interpretation of the effects of labor productivity (Y/L) and wages (w) on employment (L), is important to remember that the signs of $\beta_1 = \left(\frac{1}{\alpha + \beta - 1}\right)$ and $\beta_2 = \left(\frac{\alpha}{\alpha + \beta - 1}\right)$ and depend on the type of returns to scale.

A.2 Labor Productivity Equation

To obtain the reduced form of labor-productivity we can similarly derive the conditional demand of capital services[27],

$$\log K = \rho_0 + \rho_1 \log\left(\frac{Y}{L}\right) + \rho_2 \log w - \rho_3 \log r + \rho_4' IC + \rho_5' C + \varepsilon_K . \qquad \text{(A.8)}$$

We cannot estimate this equation directly since we do not have information on the capital stock. However, we can use equation (A.8) to obtain the reduce form of labor productivity. From equation (A.1) labor productivity can be written in levels as,

$$\left(\frac{Y}{L}\right) = TFP\, K^{\alpha}\, L^{\beta-1} \qquad \text{(A.9)}$$

and by taking logs as,

$$\log\left(\frac{Y}{L}\right) = \log TFP + \alpha \log K + (\beta - 1)\log L . \qquad \text{(A.10)}$$

Substituting logK, logL and logTFP by their corresponding equations, (A.6), (A.7) and (A.8), we get a *reduced form* equation of labor productivity in terms of the input prices (w, r), investment climate (IC) variables and other control (C) variables;

$$\log\left(\frac{Y}{L}\right) = \gamma_0 + \gamma_1 \log w + \gamma_2 \log r + \gamma_3' IC + \gamma_4' C + \varepsilon_{YL} . \qquad \text{(A.11)}$$

We know from previous empirical evidence, see for example Escribano et al. (2006), that IC variables affect labor-productivity through TFP but also through direct effects. Therfore the comments made right after equation (A.7) apply here for the interpretation of the IC effects on labor-productivity. Equation (A.11) is the *labor productivity equation* that we estimate in Table C.I of Appendix D.

A.3 Economic Interpretation of Individual Coefficients of the Employment Demand Conditional on the Level of Sales.

We will interpret how changes in TFP affect labor demand and labor-productivity.

[27] An equation similar to (A.7) and (A.8) could be derived for the demand of intermediate materials (M), when we have three inputs (L, K, and M) or when the dependent variable (Y) in the empirical application is sales and not the value added. The procedure to obtain the corresponding reduced form of labor productivity is similar and only complicates the algebra.

Consider the *conditional labor demand* (A.3) but in rates of growth,

$$\Delta\log L=\left(\frac{1}{\alpha+\beta}\right)\Delta\log Y-\left(\frac{1}{\alpha+\beta}\right)\Delta\log TFP-\left(\frac{\alpha}{\alpha+\beta}\right)\Delta\log w+\left(\frac{\alpha}{\alpha+\beta}\right)\Delta\log r. \quad \text{(A.12)}$$

Notice that since both $\alpha\ and\ \beta$ are positive, we get the expected signs of the coefficients; positive in ΔlogY and Δlog r and negative in ΔlogTFP and Δlog w.

This labor demand is derived from two equations (A.10) and (A.2). In terms of the rates of growth are,

$$\Delta\log\left(\frac{Y}{L}\right)=\Delta\log TFP+\alpha\Delta\log\left(\frac{K}{L}\right)+(\alpha+\beta-1)\Delta\log L \quad \text{(A.13a)}$$

$$\Delta\log\left(\frac{K}{L}\right)=\Delta\log\frac{\alpha}{\beta}\left(\frac{w}{r}\right)=\Delta\log\left(\frac{w}{r}\right). \quad \text{(A.13b)}$$

Keeping the input prices (w and r) constant, along the optimal path the capital labor ratio must be constant, $\Delta\log\left(\frac{K}{L}\right)$=0, and we can simplify (A.13a) and (A.13b) as,

$$\Delta\log Y=\Delta\log TFP+(\alpha+\beta)\Delta\log L\ . \quad \text{(A.14a)}$$

Suppose now that there is an improvement in TFP due to a better investment climate (IC), keeping the rest of the variables constant (caeteris-paribus).

Question 1: What is the expected *effect on employment demand* (L) of an *improvement in TFP*, keeping the level of production and the inputs prices constant?

i) Under decreasing returns to scale (DRS), $(\alpha+\beta)$<1, the reduction in labor demand will be larger that the increase in TFP.

ii) Under increasing returns to scale (IRS), $(\alpha+\beta)$>1, the reduction in labor demand will be smaller than the increase in TFP.

The reason is the following; from equations (A.12) or (A.14a) the effect of an increase in TFP, keeping the rest of the variables constant ($\Delta\log Y=0,\ \Delta\log w=0\ and\ \Delta\log r=0$), is that employment will decrease by $\Delta\log L=-\left(\frac{1}{\alpha+\beta}\right)\Delta\log TFP$.

Question 2: What is the expected *effect on labor-productivity* of an *improvement in TFP*, keeping the level of production and the input prices constant?

i) Under decreasing returns to scale (DRS), $(\alpha+\beta)$<1, the increase in labor-productivity will be larger that the increase in TFP.

ii) Under increasing returns to scale (IRS), $(\alpha+\beta)$>1, the increase in labor-productivity will be smaller that the increase in TFP.

Since $\Delta \log L = -\left(\frac{1}{\alpha+\beta}\right)\Delta \log TFP$, under the caeteris-paribus condition that $\Delta \log Y = 0, \Delta \log w = 0 \; and \; \Delta \log r = 0$ and equation (A.14b)

$$\Delta \log \left(\frac{Y}{L}\right) = \Delta \log TFP + (\alpha + \beta - 1)\Delta \log L \qquad \text{(A.14b)}$$

both changes in TFP and employment (L) will affect the changes in labor-productivity and

$$\Delta \log \left(\frac{Y}{L}\right) = \Delta \log TFP - \left(\frac{(\alpha + \beta - 1)}{\alpha + \beta}\right)\Delta \log TFP = \left(\frac{1}{\alpha + \beta}\right)\Delta \log TFP \; .$$

Question 3: What is the expected *effect on employment demand* (L) of a *reduction in wages*, keeping the level of production (Y), TFP and the rental price of capital (r) constant?

Under decreasing returns to scale (DRS), $(\alpha+\beta)$<1, the increase in labor demand due to the reduction in wages will be larger than under increasing returns to scale (IRS).

$$\Delta \log Y = \alpha\Delta \log w + (\alpha + \beta)\Delta \log L \qquad \text{(A.14c)}$$

From equations (A.12) or (A.14c) the effect of a decrease in wages, keeping the rest of the variables constant ($\Delta \log Y = 0, \Delta \log TFP = 0 \; and \; \Delta \log r = 0$), is that employment will increase by;

$$\Delta \log L = -\left(\frac{\alpha}{\alpha+\beta}\right)\Delta \log w \, .$$

Question 4: What is the expected *effect on labor-productivity* of a *reduction in wages*, keeping the level of production (Y), TFP and the rental price of capital (r) constant?

The reduction in wages will create an increase in employment demand and a reduction in labor-productivity of lower magnitude (α<1) than the decrease in wages. Therefore, the coefficient of wages in the labor-productivity equation should be positive, independently of the type of returns to scale, see equation (A.11).

From equations (A.13a) and (A.13b) and from the implied increase in L derived in question 3, we get

$$\Delta \log\left(\frac{Y}{L}\right) = \alpha\Delta \log w - (\alpha + \beta - 1)\left(\frac{\alpha}{\alpha + \beta}\right)\Delta \log w = \alpha\Delta \log w$$.

A.4 Economic Interpretation of Individual Coefficients of the Employment Demand Conditional on Labor-Productivity.

Taking first differences in equation (A.4) we get,

$$\Delta\log L=\left(\frac{1}{\alpha+\beta-1}\right)\Delta\log\left(\frac{Y}{L}\right)-\left(\frac{1}{\alpha+\beta-1}\right)\Delta\log TFP-\left(\frac{\alpha}{\alpha+\beta-1}\right)\Delta\log w+\left(\frac{\alpha}{\alpha+\beta-1}\right)\Delta\log r. \quad \text{(A.15)}$$

From equation (A.15) it is clear that the coefficients of the conditional labor demand, expressed in terms of labor-productivity (Y/L), have signs that depend on type of returns to scale ($\alpha+\beta-1$).

Question 5: What is the expected effect on labor demand of an improvement in TFP, keeping labor-productivity and the input prices constant, $\Delta\log\left(\frac{Y}{L}\right)=0, \Delta\log w=0 \; and \; \Delta\log r=0$?

i) Under decreasing returns to scale (DRS), ($\alpha+\beta-1$)<0, the improvement in TFP will *produce a positive effect on employment*, $\Delta\log L$ >0 . The intuition is the following; the initial increase in the production level (sales) due to the increase in TFP, requires an increase in employment to offset the corresponding increase in productivity, see equation (A.16). This increase in employment is larger than the initial reduction in employment created by the improvement in TFP, see question 1.

ii) Under increasing returns to scale (IRS), ($\alpha+\beta-1$)>0, the improvement in TFP will *produce a negative effect on employment*, $\Delta\log L$ <0. The intuition is the following; the initial increase in the production level (sales) due to the increase in TFP, requires a decrease in employment to offset the corresponding increase in production (Y) and in productivity, see equation (A.16).

Under the condition that labor-productivity and the input prices are constant, $\Delta\log\left(\frac{Y}{L}\right)=0, \Delta\log w=0 \; and \; \Delta\log r=0$, form equations (A.13a) and (A.13b) we have that,

$$0 = \Delta \log TFP + (\alpha + \beta - 1)\Delta \log L \quad \text{(A.16)}$$

or equivalently, from (A.15) the effect of an increase in TFP is $\Delta\log L=-\left(\frac{1}{\alpha+\beta-1}\right)\Delta\log TFP$.

Question 6: What is the expected effect on employment demand of *an increase in labor-productivity,* keeping TFP and the input prices constant, $\Delta \log TFP = 0, \Delta \log w = 0 \; and \; \Delta \log r = 0$?

i) Under decreasing returns to scale (DRS), $(\alpha + \beta - 1)$<0, the increases in labor-productivity is due to the fact that the firm is using less employment, $\Delta \log L$<0, see equation (A.17).

ii) Under increasing returns to scale (IRS), $(\alpha + \beta - 1)$>0, the increases in labor-productivity must be due to the fact that the firm is using more employment and producing proportionally more output, see equation (A.17), creating a positive effect on employment, $\Delta \log L$ >0.

Under the conditions that TFP and the input prices are constant, $\Delta \log TFP = 0, \Delta \log w = 0 \; and \; \Delta \log r = 0$, from equations (A.13a) and (A.13b) we have that,

$$\Delta \log \left(\frac{Y}{L} \right) = (\alpha + \beta - 1) \Delta \log L = 0 \; . \tag{A.17}$$

From (A.15) or (A.17) the labor-productivity effect is

$$\Delta \log L = \left(\frac{1}{\alpha + \beta - 1} \right) \Delta \log \left(\frac{Y}{L} \right)$$

Question 7: What is the expected *effect on employment* (L) of a *reduction in wages*, keeping the level of labor-productivity constant (Y/L), TFP and the rental price of capital (r) constant?

i) Under decreasing returns to scale (DRS), $(\alpha + \beta)$<1, there will be a *decrease in labor demand due to a reduction in wages.* The initial decrease in wages creates a reduction in labor-productivity due to an increase of employment which creates a further reduction in labor productivity. In order to keep the labor-productivity constant there should be a decrease in employment larger than the initial increase.

ii) Under increasing returns to scale (IRS), $(\alpha + \beta)$>1, there will be an *increase in labor demand due to the reduction in wages.* The initial decrease in wages creates a reduction in labor-productivity due to an increase of employment which creates a larger increase in production to keep labor-productivity constant.

The effect of a reduction in wages (w), keeping the rest of the variables constant ($\Delta \log \left(\frac{Y}{L} \right) = 0, \Delta \log TFP = 0 \; and \; \Delta \log r = 0$) requires,

$$\Delta \log \left(\frac{Y}{L} \right) = \alpha \Delta \log w + (\alpha + \beta - 1) \Delta \log L = 0 \tag{A.18}$$

and therefore, from equation (A.15) or (A.18) $\Delta \log L = -\left(\frac{\alpha}{\alpha + \beta - 1} \right) \Delta \log w$.

Appendix B: Data Transformations

Sampling Frame

The sampling frame for the Retail ICS was the list of the retail stores interviewed by AC Nielsen for inventory verification on behalf of distributions of branded goods. This list covered 1433 retail stores in 41 cities across India for three industries: fast moving consumer goods (FMCG), modern format stores and consumer durable stores.

Labor Cost

Step 1: *Calculate the number of workers*. Number of workers is defined as total number of full time workers (permanent or temporary). Due to the large amount of zeros in the original number of permanent workers variable, in order to solve the problem, we transform the original sample with missing values by adding 1.25 workers to all the stores with less than 3.75 workers. Therefore, all stores included in this replacement process are either *micro or small stores* since maximum number of employees becomes (1.25+3.75) = 5 employees. In the case of temporary workers, we multiply this number by the average length of the contract in months divided by 12.

Step 2: Unit labor cost. Total annual cost of labor divided by the number of workers.

Step 3: *Missing values.* We began by stratifying original sample into sub-groups (by type of store and region, 12 sub-groups) in order to compute the median of unit labor cost for each group them, the missing values of labor cost are replaced by these medians multiplied by the number of workers.

Step 4: *Outliers in labor cost*. Final step slightly modify the sample we obtained in step 3; it simply consists in excluding the outliers which were defined as those observations with ratios of labor cost to sales greater than one.

Total Rental Cost of Capital (Interpolation)

Step 1: *Estimation of the total rental cost of capital*. With the available observations, we run an OLS regression of log of the total rental cost of capital on the log of *selling area of the store in square yards*; this regression includes constant term and industry (type of store) and store size dummies.

Table B.1: Results of the OLS Regression of Log of Rental Cost of Capital

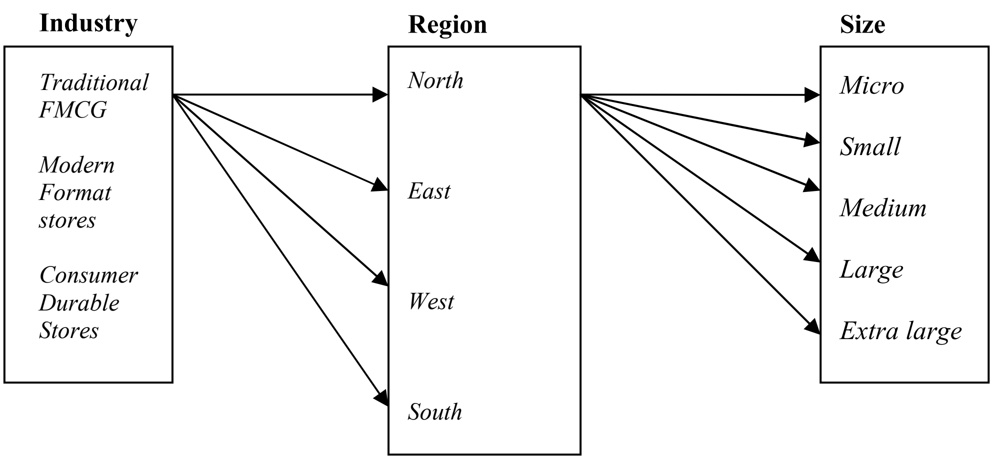
Explanatory Variables	Coefficient
Log of area	0.362***
Micro	-1.462***
Small	-0.827
Medium	-0.389
Extra-large	-0.030
Traditional FMCG	-0.463**
Consumer Durable Stores	-0.231
Constant	6.141***
Observations	*1110*
R-squared	*0.26*

* significant at 10%; ** significant at 5%; *** significant at 1%

Step 2: *Missing values.* The missing values in the total rental cost of capital are replaced by their predictions obtained from the regression of step1 (interpolation).

Step 3: *Outliers in the total rental cost of capital*. Final step slightly modify the sample we obtained in step 2; it simply consists in excluding the outliers which were defined as those observations with ratios of rental cost of capital to sales greater than one.

Figure B.1: Stratification Process:



* Variable "Size" used here is based in the number of employees. *Micro* firms are those with less than 2 employees; *small* firms are in between 2 and 5 employees; *medium* firms are within 5 and 15 employees; *large* firms are in between 15 and 50 employees; finally, *extra large* firms are those with more than 50 employees.

Table B.2: Total Number of Observations, Missing Values and Zeros for Labor Cost and Rental Cost of Capital Variables in the Original Sample Before and After Dropping Outliers.

	Labor cost	Rental cost of capital
Maximum number of observations.	***1948***	***1948***
(1) Missing Values	26	32
(2) Zeros	842	812
(3) Observations not available: (1)+(2)	868	844
Available observations including outliers	1080	1104
Final number of observations after correcting for outliers and missing observations.[1]	**1927**	**1897**

1See the appendix for a description of the methodology used to deal with outliers and to replace missing values and zeros.

Appendix C: Definitions

I Production Function Variables[28]

Sales: Used as the measure of output for the production function estimation. Sales are defined as the store sales of last complete fiscal year's, total annual sales (2005-06).

Employment: Total number of full time workers (permanent or temporary).

Total hours worked per year: Total number of employees multiplied by the average hours worked per year.

Total rental cost of capital: The rental cost of capital is defined as total annual cost of rental land/buildings, equipment and furniture.

Rental cost of capital: Total rental cost of capital per square yard.

Labor cost: Total expenditures on personnel.

Wages: Labor cost per worker.

II Dependent Variable in Equation Regressions and Linear Probability Models

Demand for Labor: Total number of permanent and temporal workers.

III General Information at Plant Level

Sector classification: a) Traditional Fast Moving Consumer Goods (FMCG); b) Modern Format Stores; c) Consumer Durable Stores.

Regional classification: a) North; b) East; c) West; d) South.

Size classification: a) micro firms (< 2 employees); b) small firms (>=2 & <5); c) medium firms (>=5 & <15); d) large firms (>=15 & <50); e) extra-large firms (>=50).

[28] All series figure in US dollars, data obtained from The World Bank.

Table A.I: Investment climate (IC) and control (C) variables.

Blocks of ICAs	Name of the variable	Description of the variable	Observations (Response rate %)
Infrastructures	Dummy for power outages	Dummy variable that takes value 1 if the store has suffered any power outages in last year.	1944(99.8)
	Number of power outages	Number of power outages suffered by a store.	1611(82.7)
	Average duration of power outages	Average duration of power outages suffered by the store in hours.	1611(82.7)
	Losses due to power outages	Value of the losses due to power outages as a percentage of sales (conditional on the store reporting power outages).	1326(68.1)
	Days of inventory	Days of inventory of main sales item.	1948(100)
Red tape, informality and others	Dummy for criminal attempts	Dummy variable that takes value 1 if the store suffered any criminal attempt during last year.	1944(99.8)
	Losses due to criminal activity	Value of losses due to criminal activity.	1834(94.1)
	Dummy for security	Dummy variable that takes value 1 if the store has paid for security.	1947(99.9)
	Manager's time spent in bureaucratic issues	Percentage of managers' time spent in dealing with bureaucratic issues.	1948(100)
	Dummy for payments to deal with bureaucratic issues	Dummy that takes value 1 if store in the main sector occasionally need to give gifts or make informal payments to public officers in order to "get things done" with regard to customs, taxes, licenses, legislations, services, etc.	1418(72.8)
	Sales declared to taxes	Percentage of total sales declared to taxes.	1669(85.7)
	Labor costs declared	Percentage of workforce declared to taxes.	1632(83.8)
	Dummy for agency visit	Dummy variable that takes value 1 if the store has been visited by any centre, state or local agency.	1937(99.4)
	Overdue payments	Number of days to resolve overdue payments.	1181(60.6)
	Dummy for third party	Dummy variable that takes value 1 if the store had to engage third party to resolve a dispute with clients over payments.	1947(99.9)
	Sales never repaid	Percentage of monthly total sales to private customers that were never repaid.	1181(60.6)

Table A.II: Investment climate (IC) and control (C) variables.

Blocks of ICAs	Name of the variable	Description of the variable	Observations (Response rate %)
Finance and Accountability	Dummy for credit line	Dummy variable that takes value 1 if the store reports that it has a credit line.	1928(99)
	Dummy for loan	Dummy variable that takes value 1 if the store reports that it has a bank loan.	1921(98.6)
	Dummy for external audit	Dummy variable that takes value 1 if store's annual statements are engaged in a process of external audit.	1914(98.3)
	Financing - Internal founds	Percentage of financing that comes from internal founds.	1948(100)
	Financing - State-owned banks	Percentage of financing that comes from state-owned banks.	1948(100)
	Financing - Private commercial banks	Percentage of financing that comes from private commercial banks.	1948(100)
	Financing - Non-banks financial institutions	Percentage of financing that comes from non-banks financial institutions.	1948(100)
	Financing - Family/friends loans	Percentage of financing that comes from family/friends loans.	1948(100)
	Financing - Informal sources	Percentage of financing that comes from informal sources.	1948(100)
	Financing - Credit from suppliers	Percentage of financing that comes from credits from suppliers.	1948(100)
	Dummy for rent land	Dummy variable that takes value 1 if the store rents almost all its lands.	1948(100)
	Dummy for rent buildings	Dummy variable that takes value 1 if the store rents almost all its buildings.	1948(100)
	Dummy for current account	Dummy variable that takes value 1 if the firm has a current or saving account	1940(99.6)
Quality, Innovation and Labor Skills	Experience of the manager	Number of years of experience of the manager.	1948(100)
	Dummy for computer	Dummy variable that takes value 1 if the store uses computer.	1948(100)
Other Control Variables	Dummy for individual proprietorship	Dummy variable that takes value 1 if the owner of the store is a single person.	1948(100)
	Dummy for part of a larger firm	Dummy variable that takes value 1 if the store is a part of a larger firm.	1927(98.9)
	Age of the firm	Difference between the year that the plant started operations and current year.	1948(100)

Appendix D: Tables and figures

Table B.1: Representativeness of production function variables before and after cleaning missing values and outliers; by industry and region.

Region		North		East		West		South		Total	
Sector		#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.
Traditional FMCG	Original Sample	330	16.94	207	10.63	369	18.94	338	17.35	1244	63.86
	Without replacing	53	8.18	32	4.94	67	10.34	152	23.46	304	46.91
	With replacing	322	16.97	204	10.75	354	18.66	335	17.66	1215	64.05
Modern Format Stores	Original Sample	30	1.54	17	0.87	71	3.64	71	3.64	189	9.70
	Without replacing	11	1.70	14	2.16	31	4.78	55	8.49	111	17.13
	With replacing	27	1.42	16	0.84	66	3.48	68	3.58	177	9.33
Consumer Durable Stores	Original Sample	152	7.80	79	4.06	135	6.93	149	7.65	515	26.44
	Without replacing	40	6.17	35	6.17	51	7.87	107	16.51	233	35.96
	With replacing	149	7.85	79	4.16	130	6.85	147	7.75	505	26.62
Total	Original Sample	512	26.28	303	15.55	575	29.52	558	28.64	1948	100.00
	Without replacing	104	16.05	81	12.50	149	22.99	314	48.46	648	100.00
	With replacing	498	26.25	299	15.76	550	28.99	550	28.99	1897	100.00

Source: Authors' calculations with India ICS data.

Table B.2: Percentage of observations lost due to missing values; by industry and region.

Region		North		East		West		South		Total	
Sector		#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.	#Obs.	Perc.
Traditional FMCG	Original Sample	330		207		369		338		1244	
	Without replacing	53	83.94	32	84.54	67	81.84	152	55.03	304	75.56
	With replacing	322	2.42	204	1.45	354	4.07	335	0.89	1215	2.33
Modern Format Stores	Original Sample	30		17		71		71		189	
	Without replacing	11	63.33	14	17.65	31	56.34	55	22.54	111	41.27
	With replacing	27	10.00	16	5.88	66	7.04	68	4.23	177	6.35
Consumer Durable Stores	Original Sample	152		79		135		149		515	
	Without replacing	40	73.68	35	55.70	51	62.22	107	62.22	233	54.76
	With replacing	149	1.97	79	0.00	130	3.70	147	1.34	505	1.94
Total	Original Sample	512		303		575		558		1948	
	Without replacing	104	79.69	81	73.27	149	74.09	314	43.73	648	66.74
	With replacing	498	2.73	299	1.32	550	4.35	550	1.43	1897	2.62

Source: Authors' calculations with India ICS data.

Table C.I: IC Elasticities and Semi-elasticities with respect to Log-Labor-Productivity.

	Retail Sector		Traditional FMCG		Modern Format Stores		Consumer Durable Stores	
Explanatory Variables	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Log of wages	0.342***	85.8	0.357***	103.5	0.172**	59.1	0.446***	94.4
Log of rental cost of capital	0.092***	26.1	0.091***	29.0	0.044	19.2	0.105**	25.8
Infrastructures								
Av. Duration of power outages (b)	-0.177***	-3.8	n.s.		-0.420***	-8.2	n.s.	
Losses due to power outages (a)	n.s.		-0.051**	-6.9	n.s.		-0.056***	-9.7
Days of inventory	0.061**	5.3	0.049*	4.3	n.s.		n.s.	
Red tape, informailty and others								
Dummy for criminal attempts	-0.198**	-0.7	n.s.		n.s.		n.s.	
Dummy for security	0.203***	2.4	0.159**	1.5	0.276*	8.4	0.285**	4.1
Dummy for store visited by any agency	0.224***	4.0	0.244***	4.8	n.s.		0.262**	4.0
Dummy for third party arbitrage	0.267**	0.5	0.444***	0.4	0.934**	1.2	n.s.	
Finance and Accountability								
Dummy for credit line	0.167**	0.9	0.231**	1.1	n.s.		0.277*	1.9
Dummy for current account	0.334***	9.6	0.382***	11.0	n.s.		0.352***	10.3
Dummy for external audit	0.249***	3.3	0.161*	1.6	n.s.		0.362***	5.7
Financing-Internal funds	n.s.		n.s.		n.s.		0.004**	10.2
Financing-Family sources	-0.003***	-0.9	-0.003***	-1.3	-0.013***	-1.9	n.s.	
Financing-Informal sources	0.005*	0.2	0.009***	0.4	n.s.		n.s.	
Dummy for additional land	0.25*	0.3	0.547**	0.5	n.s.		n.s.	
Quality, Innovation and Labor Skills								
Experience of the manager	0.066*	6.8	n.s.		0.284***	33.5	0.124*	10.6
Dummy for computer	0.415***	3.0	0.699***	1.9	n.s.		n.s.	
Other Control Variables								
Dummy for part of a larger firm	n.s.		n.s.		n.s.		0.795***	28.7
Age of the firm	0.085**	8.4	0.121***	14.2	-0.126*	-14.0	n.s.	
Observations	1814		1166		175		487	
R-squared	0.35		0.23		0.5		0.31	

Notes:
*significant at 10%; ** significant at 5%; *** significant at 1% given by robust standard errors corrected for correlation between cluster (industry and region).
Each regression includes a set of industry, size and region dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
n.s.: variable not significant for this regression.
Source: Authors' calculations with India ICS data.

Table C.II: IC percentage contributions to aggregate log-productivity. Whole Retail Sector

		Aggregate log-labor-productivity	Average log-labor-productivity	Allocative efficiency
Log of wages		93.9	85.8	8.2
Log of rental cost of capital		30.1	26.1	3.9
Infrastructures	Av. Duration of power outages (b)	-3.3	-3.8	0.5
	Days of inventory	6.3	5.3	1.0
Red Tape, informality and others	Dummy for criminal attempts	-1.0	-0.7	-0.3
	Dummy for security	4.7	2.4	2.2
	Dummy for store visited by any agency	5.6	4.0	1.6
	Dummy for third party arbitrage	1.0	0.5	0.5
Finance and accountability	Dummy for credit line	1.8	0.9	0.9
	Dummy for current account	13.4	9.6	3.7
	Dummy for external audit	7.2	3.3	3.8
	Financing-Family sources	-0.6	-0.9	0.3
	Financing-Informal sources	0.2	0.2	0.0
	Dummy for additional land	0.8	0.3	0.5
Quality, innovation and labor skills	Experience of the manager	7.2	6.8	0.5
	Dummy for computer	9.2	3.0	6.2
Other control variables	Age of the firm	9.1	8.4	0.7
Total contribution of IC (demean log-labor-productivity)		61.6	39.5	22.2
Other stuff	Industry/region/size controls	-20.5	-13.5	-7.0
	Constant term	-99.2	-99.2	0.0
	Residual	34.1	0.0	34.1
Total contribution of other stuff		-85.6	-112.7	27.1
Total		100.0	38.7	61.3

NOTES:
Results from equation 2.3 .
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.III: IC percentage contributions to aggregate log-productivity. Traditional FMCG Stores

		Aggregate log-labor-productivity	Average log-labor-productivity	Allocative efficiency
Log of wages		109.6	103.5	6.1
Log of rental cost of capital		32.1	29.0	3.1
Infrastructures	Losses due to power outages (a)	-6.9	-6.9	0.0
	Days of inventory	4.8	4.3	0.4
Red Tape, informality and others	Dummy for security	2.6	1.5	1.1
	Dummy for store visited by any agency	6.8	4.8	2.0
	Dummy for third party arbitrage	0.5	0.4	0.0
Finance and accountability	Dummy for credit line	1.8	1.1	0.7
	Dummy for current account	16.3	11.0	5.2
	Dummy for external audit	3.7	1.6	2.1
	Financing-Family sources	-0.9	-1.3	0.4
	Financing-Informal sources	0.7	0.4	0.3
	Dummy for additional land	1.2	0.5	0.7
Quality, innovation and labor skills	Dummy for computer	7.6	1.9	5.7
Other control variables	Age of the firm	15.8	14.2	1.6
Total contribution of IC (demean log-labor-productivity)		53.8	33.5	20.3
Other stuff	Industry/region/size controls	-14.5	-7.7	-6.9
	Constant term	-129.0	-129.0	0.0
	Residual	48.0	0.0	48.0
Total contribution of other stuff		-95.5	-136.7	41.2
Total		100.0	29.3	70.7

NOTES:
Results from equation 2.3
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.IV: IC percentage contributions to aggregate log-productivity. Modern Format Stores

		Aggregate log-labor-productivity	Average log-labor-productivity	Allocative efficiency
Log of wages		60.3	59.1	1.2
Log of rental cost of capital		20.2	19.2	1.0
Infrastructures	Av. Duration of power outages (b)	-6.6	-8.2	1.5
Red Tape, informality and others	Dummy for security	10.9	8.4	2.5
	Dummy for third party arbitrage	2.5	1.2	1.3
Finance and accountability	Financing-Family sources	-0.8	-1.9	1.0
Quality, innovation and labor skills	Experience of the manager	35.9	33.5	2.4
Other control variables	Age of the firm	-14.9	-14.0	-0.9
Total contribution of IC (demean log-labor-productivity)		26.9	19.0	7.9
Other stuff	Industry/region/size controls	-81.3	-66.1	-15.2
	Constant term	51.5	51.5	0.0
	Residual	22.4	0.0	22.4
Total contribution of other stuff		-7.4	-14.6	7.1
Total		100.0	82.7	17.3

NOTES:
Results from equation 2.3
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table C.V: IC percentage contributions to aggregate log-productivity. Consumer Durable Stores

		Aggregate log-labor-productivity	Average log-labor-productivity	Allocative efficiency
Log of wages		101.3	94.4	6.9
Log of rental cost of capital		27.6	25.8	1.8
Infrastructures	Losses due to power outages (a)	-7.9	-9.7	1.8
Red Tape, informality and others	Dummy for security	5.6	4.1	1.5
	Dummy for store visited by any agency	4.9	4.0	0.9
Finance and accountability	Dummy for credit line	3.2	1.9	1.3
	Dummy for current account	11.9	10.3	1.7
	Dummy for external audit	8.9	5.7	3.3
	Financing-Internal founds	9.3	10.2	-0.8
Quality, innovation and labor skills	Experience of the manager	11.4	10.6	0.7
Other control variables	Dummy for part of a larger firm	27.8	28.7	-1.0
Total contribution of IC (demean log-labor-productivity)		75.2	65.8	9.3
Other stuff	Industry/region/size controls	-20.5	-13.5	-7.0
	Constant term	-99.2	-99.2	0.0
	Residual	34.1	0.0	34.1
Total contribution of other stuff		-85.6	-112.7	27.1
Total		118.5	73.3	45.2

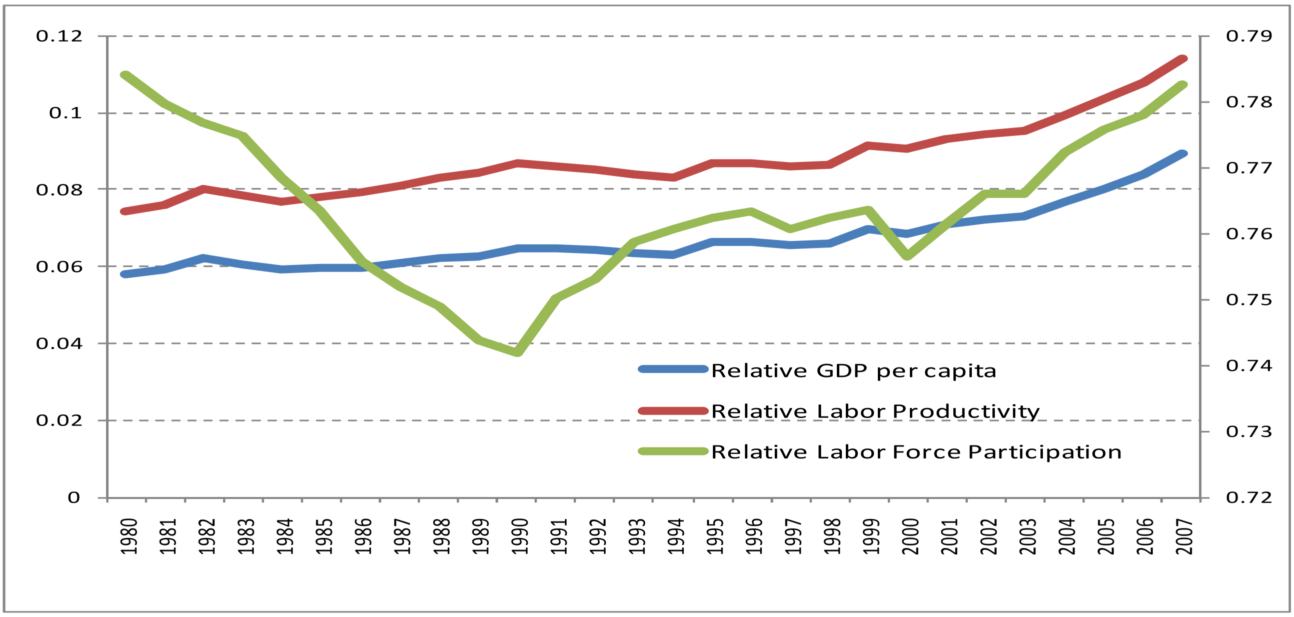
NOTES:
Results from equation 2.3
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
Source: Authors' calculations with India ICS data.

Table D: IC elasticities and semi-elasticities with respect to employment – IV Estimation.

Dependent variable: employment		**Retail Sector**		**Micro and Small Firms**	
Blocks	**Explanatory ICA variables**	**Coefficient**	**% Contrib**	**Coefficient**	**% Contrib**
Log of labor-productivity1		0.253***	37.8	0.211***	34.4
Log of wages2		-0.473*	-616.7	-0.465*	-686.3
Log of rental cost of capital		0.146***	197.9	0.082***	125.4
Infrastructures	Days of inventory	-0.049**	-20.5	-0.026*	-12.4
Red Tape, informality and others	Dummy for security	0.140**	6.0	n.s.	
	Overdue payments (a)	n.s.		-0.103**	-79.3
Finance and corporate governance	Dummy for credit line	n.s.		-0.099*	-2.3
	Dummy for current account	n.s.		0.095**	12.6
	Dummy for overdraft	0.262***	6.6	n.s.	
	Dummy for external audit	0.193***	6.4	0.213***	7.1
	Financing-Internal funds	n.s.		-0.001*	-19.5
Quality, innovation and labor skills	Dummy for computer	0.499***	6.9	0.183**	2.1
Other control variables	Age of the firm	0.092***	46.9	n.s.	
Instruments evaluation	First stage R-squared: productivity2	0.244		0.253	
	Partial R-squared: productivity3	0.025		0.024	
	Partial R-squared F test (p-value):productivity4	0.000		0.000	
	First stage R-squared: wages2	0.201		0.135	
	Partial R-squared: wages3	0.014		0.008	
	Partial R-squared F test (p-value): wages4	0.001		0.089	
	Hansen test (p-value)[5]	0.273		0.787	
	Observations	1826		1560	

NOTES:
* significant at 10%; ** significant at 5%; *** significant at 1% (robust standard errors corrected for clustering by industry and region). Each regression includes a set of industry and region dummies and a constant term.
(a) Variables instrumented with the industry-region-size average.
(b) Variables approximated with a proxy (only missing values replaced by the industry-region-size average).
[1] Labor-productivity and real wages are endogenous and the list of variables used as excluded instruments comes from the list of explanatory variables from their corresponding equations.
[2] First stage R-squared from the regression of productivity on both the included and the excluded instruments.
[3] The partial R-squared measures the squared partial correlation between the excluded instruments and the productivity.
[4] F-test of joint significance of the excluded instruments that corresponds to the partial R-squared.
[5] The Hansen test is a test of overidentifying restrictions. The null hypothesis is that the instruments are valid instruments, that is, uncorrelated with the error term, and therefore the excluded instruments are correctly excluded from the estimated equation.
n.s.: variable not significant for this equation.
Source: Authors' calculations with India ICS data.

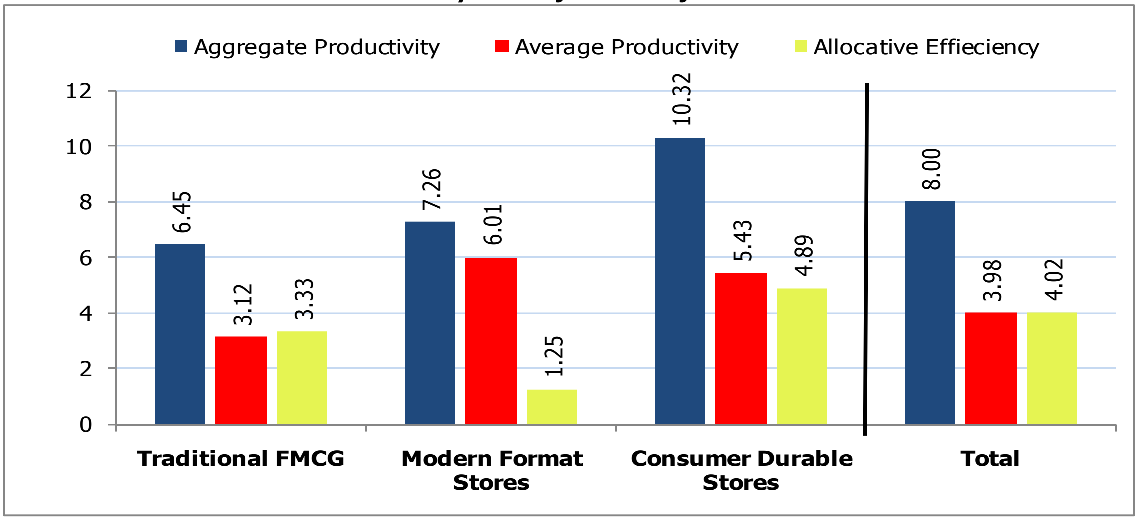
Figure 8: Decomposition of GDP gap between India and USA, 1980/2007



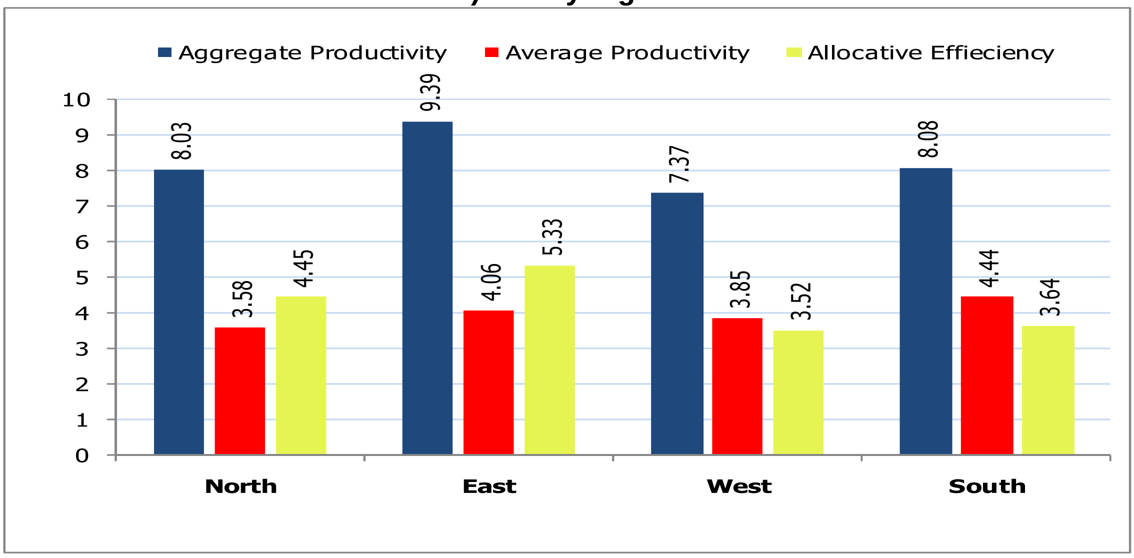
Source: Authors `Calculations with Penn World Table Version 6.3, Center for International Comparisons at the University of Pennsylvania.

Figure 9: Olley and Pakes decomposition in levels.

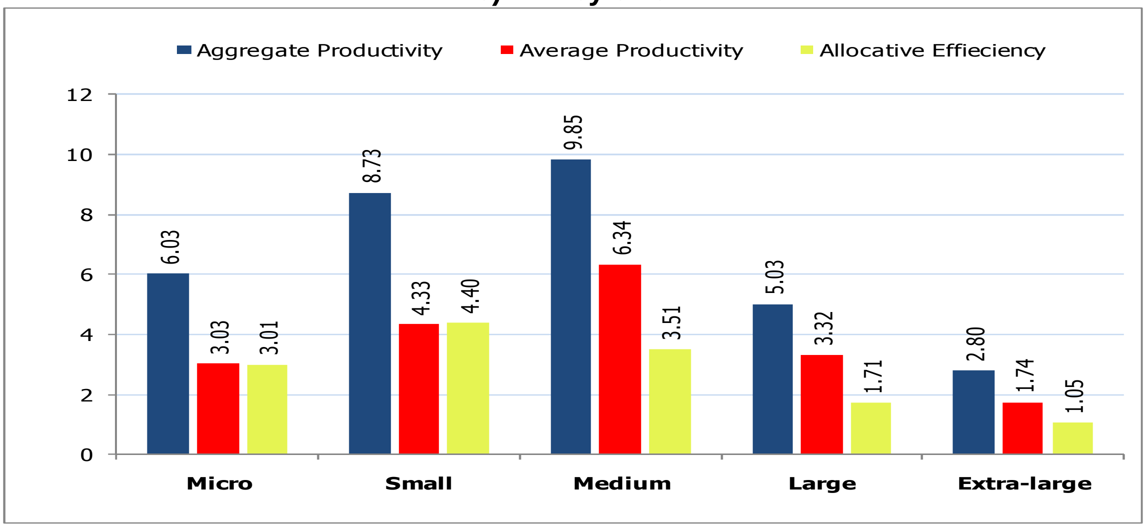
a) by industry



b) by region



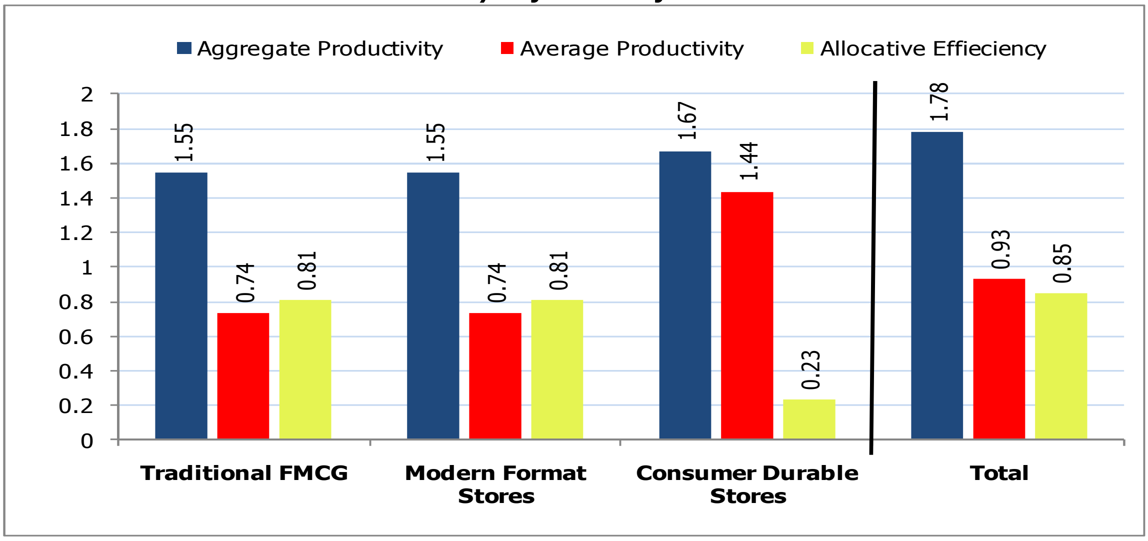
c) by size



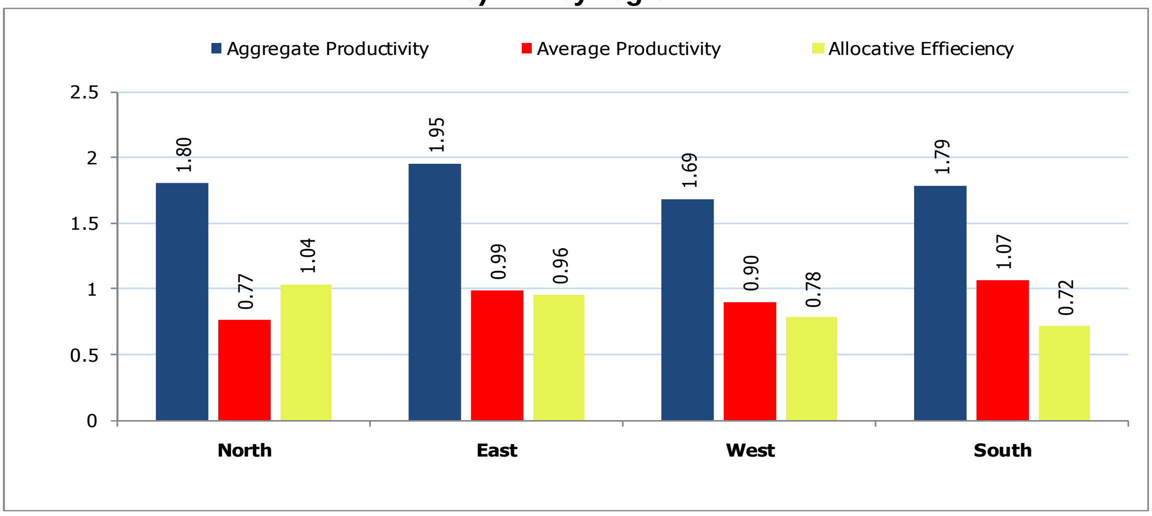
Note: Olley and Pakes decomposition in levels according to equation 2.3.
Source: Authors' calculations with India ICS data.

Figure 3: *Mixed* Olley and Pakes decomposition.

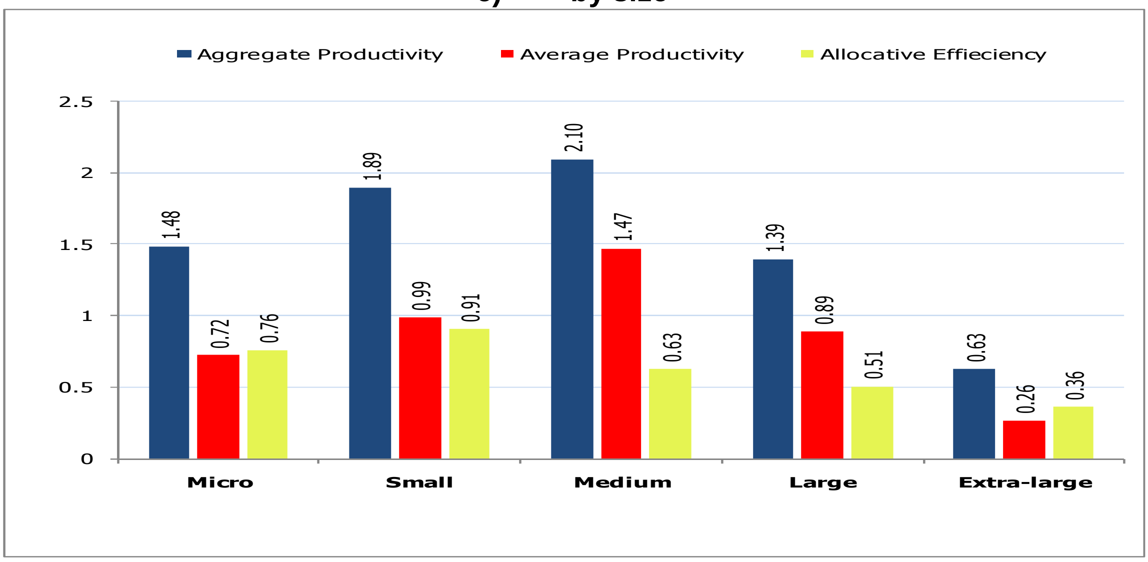
a) by industry



b) by region

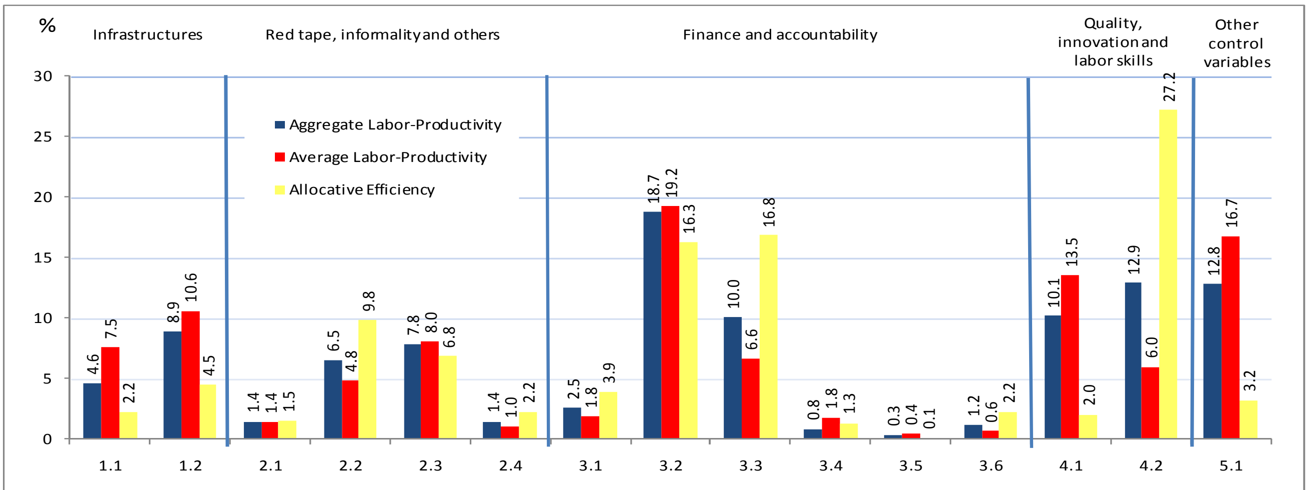


c) by size



Note: Mixed Olley and Pakes decomposition according to equation 2.3.
Source: Authors' calculations with India ICS data.

Figure 4.1: IC Percentage contributions to aggregate log-labor-productivity. Whole Retail Sector



Legend:

1. Infrastructures
1.1 Av. Duration of power outages.
1.2 Days of inventory.

2. Red tape, informality and others
2.1 Dummy for criminal attempts
2.2 Dummy for security
2.3 Dummy for store visited by any agency
2.4 Dummy for third party arbitrage

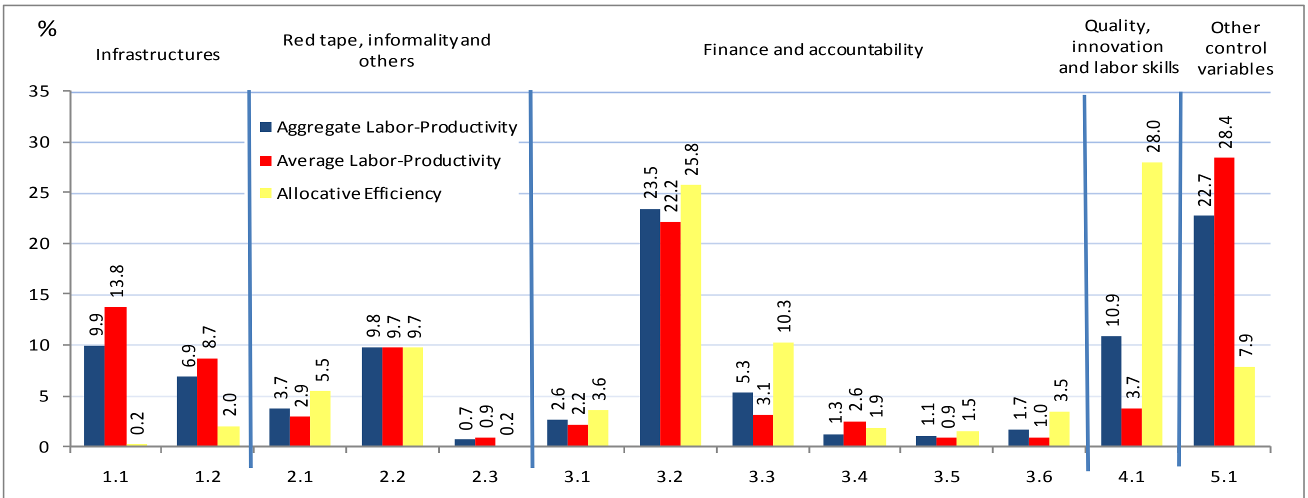
3. Finance and accountability
3.1 Dummy for credit line
3.2 Dummy for current account
3.3 Dummy for external audit
3.4 Financing-Family sources
3.5 Financing-Informal sources
3.6 Dummy for additional land

4. Quality, innovation and labor skills
4.1 Experience of the manager
4.2 Dummy for computer.

5. Other control variables
5.1 Age of the firm.

Note: Contributions computed according to section 2.
Source: Authors' calculations with India ICS data.

Figure 4.2: IC Percentage contributions to aggregate log-labor-productivity. Traditional FMCG



Legend:

1. Infrastructures
1.1 Losses due to power outages.
1.2 Days of inventory.

2. Red tape, informality and others
2.1 Dummy for security
2.2 Dummy for store visited by any agency
2.3 Dummy for third party arbitrage

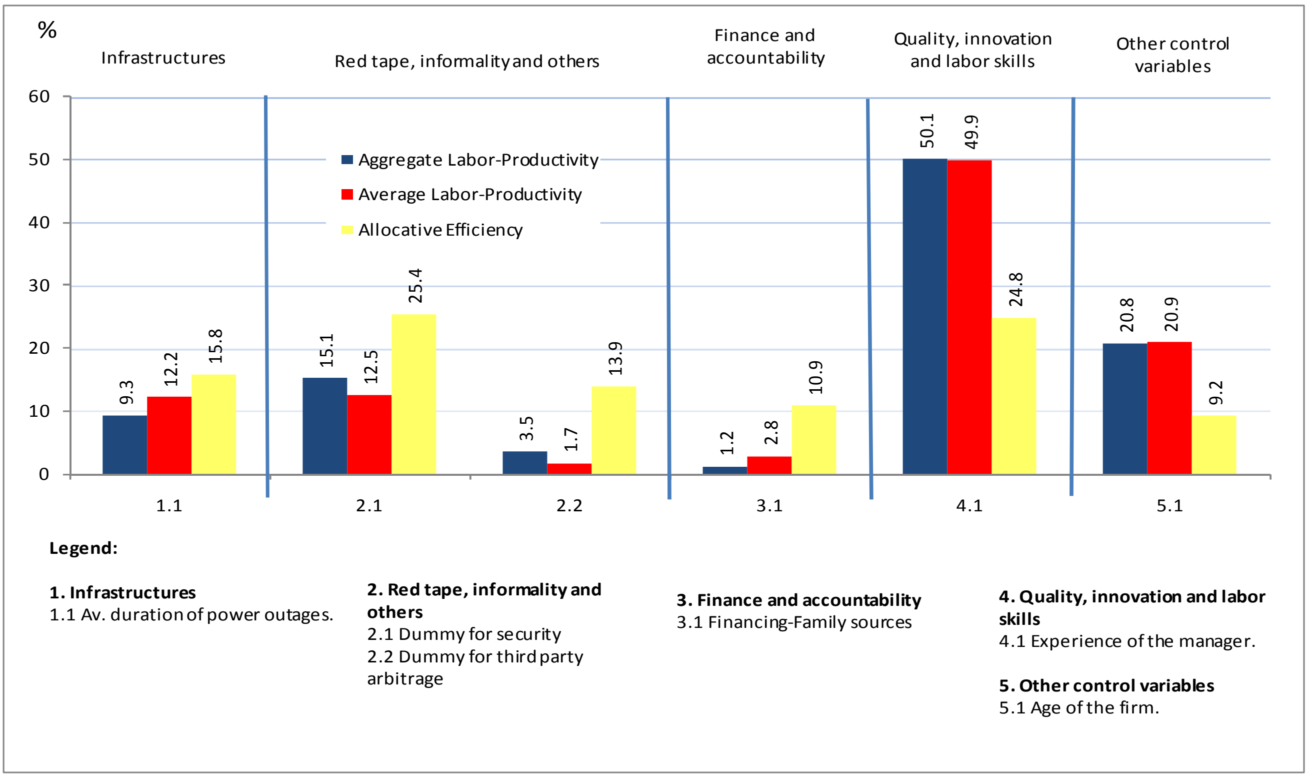
3. Finance and accountability
3.1 Dummy for credit line
3.2 Dummy for current account
3.3 Dummy for external audit
3.4 Financing-Family sources
3.5 Financing-Informal sources
3.6 Dummy for additional land

4. Quality, innovation and labor skills
4.1 Dummy for computer.

5. Other control variables
5.1 Age of the firm.

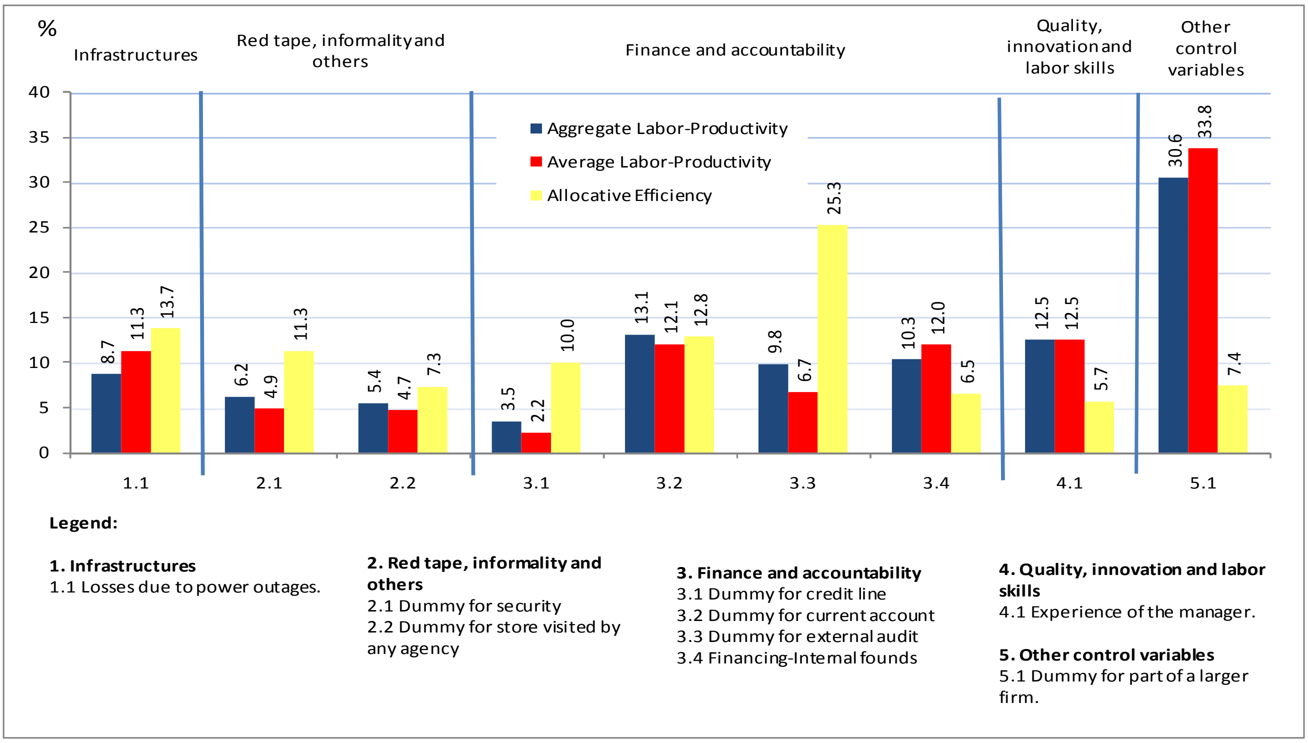
Note: Contributions computed according to section 2.
Source: Authors' calculations with India ICS data.

Figure 4.3: IC Percentage contributions to aggregate log-labor-productivity. Modern Format Stores



Note: Contributions computed according to section 2.
Source: Authors' calculations with India ICS data.

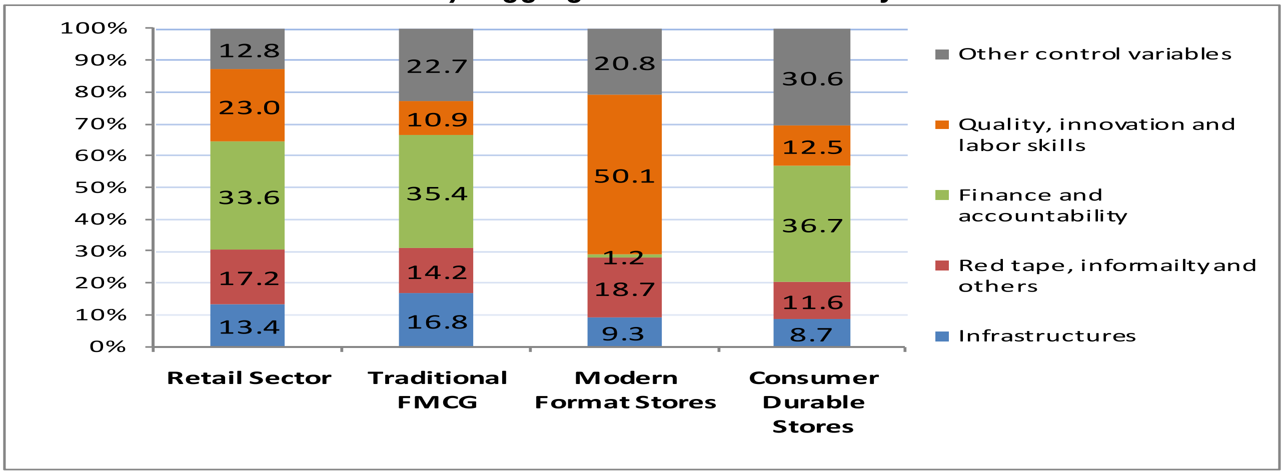
Figure 4.4: IC Percentage contributions to aggregate log-labor-productivity. Consumer Durable Stores



Note: Contributions computed according to section 2.
Source: Authors' calculations with India ICS data.

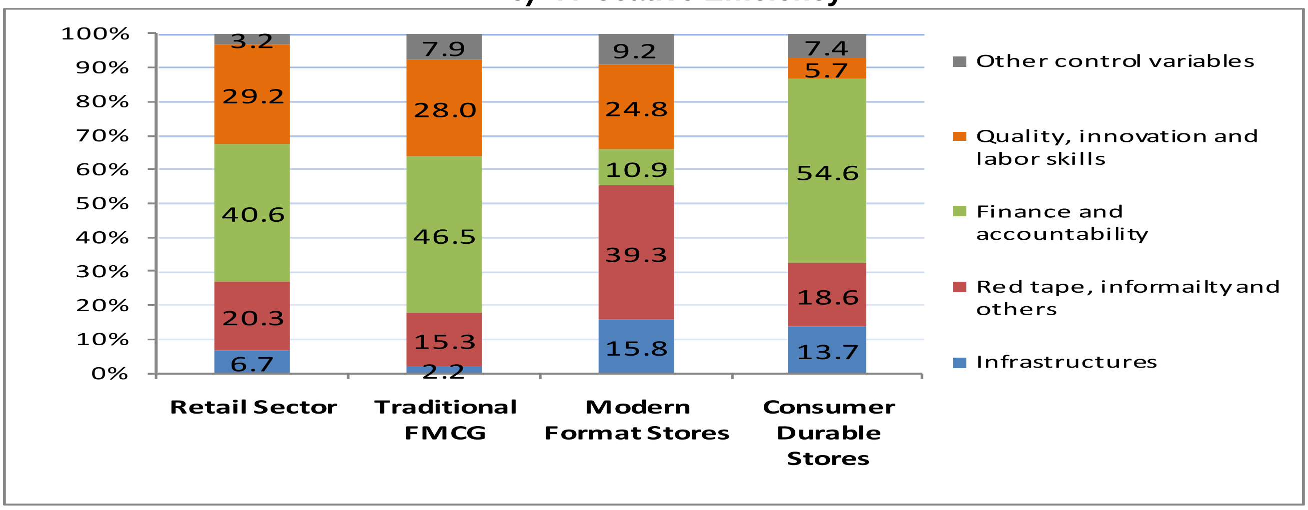
Figure 4.5: Weight of each block of IC variables on Olley and Pakes decomposition.

a) Aggregate Labor-Productivity



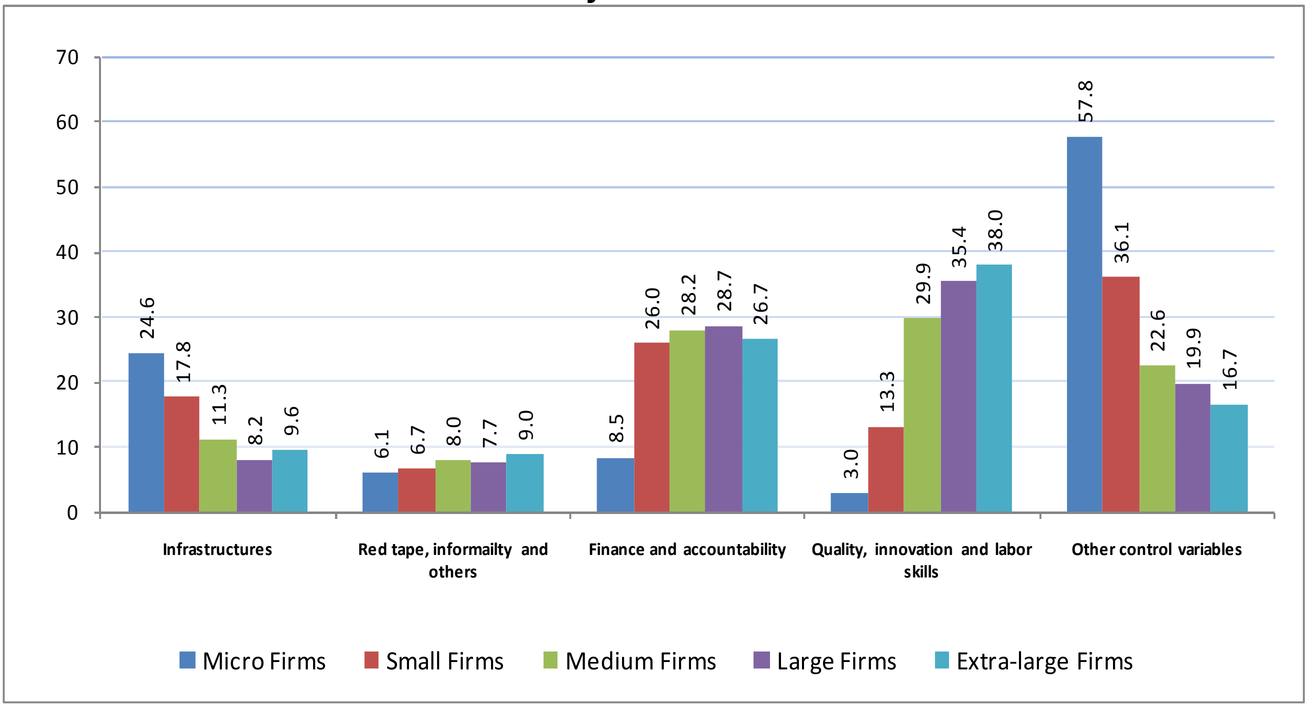
b) Average Labor-Productivity

c) Allocative Efficiency



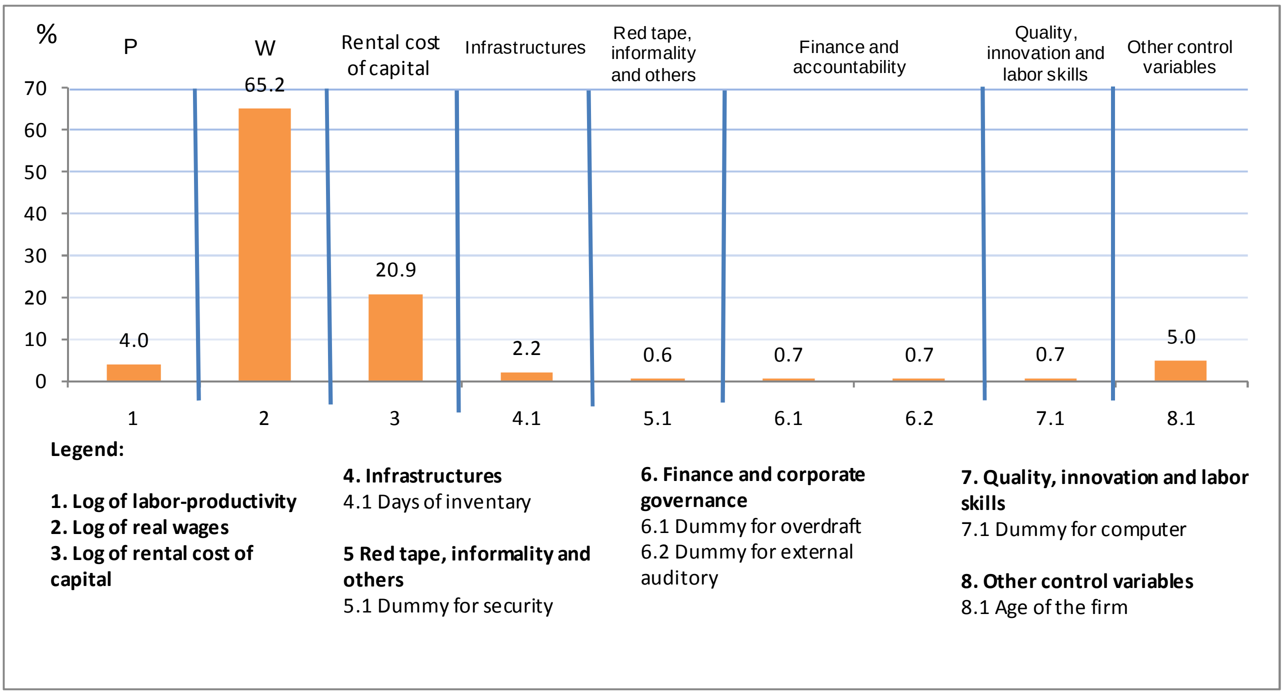
Note: The weight of each block or group of IC variables from contributions comes from Figures 4.1-4.4. We take the percentage contributions of Tables C.II-C.V in absolute value and we compute the relative weight of each block.
Source: Authors' calculations with India ICS data.

Figure 4.6: IC Percentage Absolute Contribution on Average Labor-productivity by Size



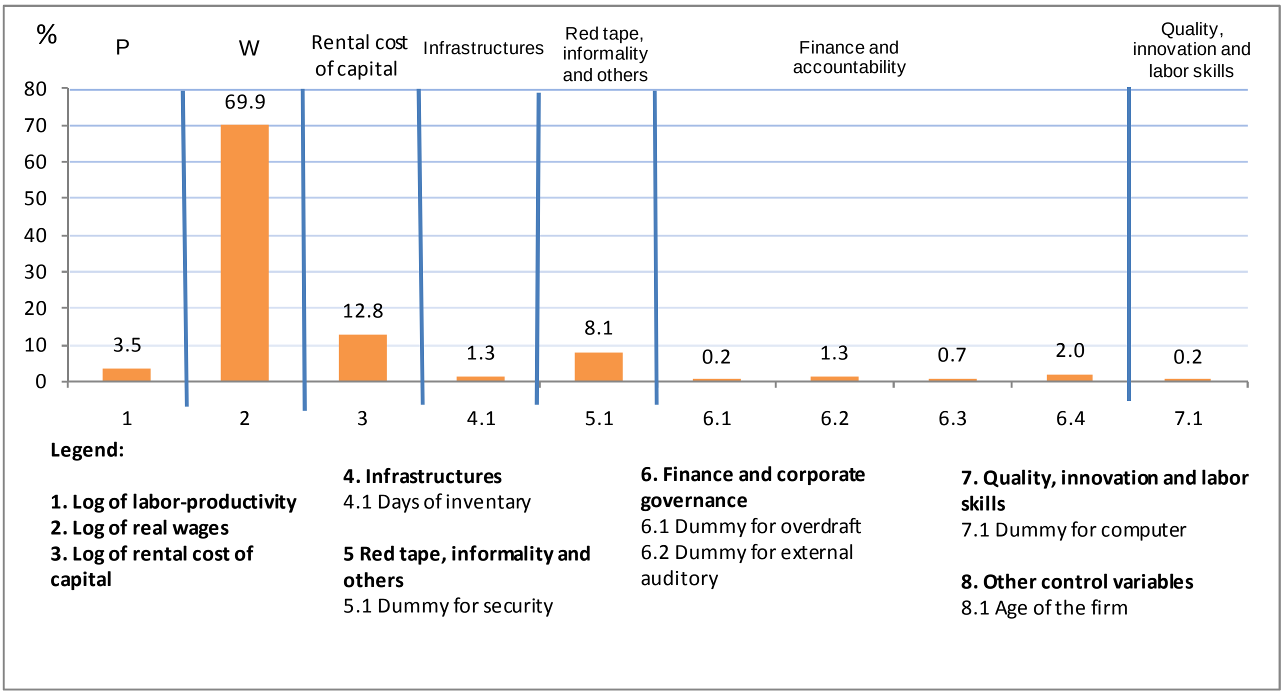
Note: Contributions computed according to section 2.
Source: Author's calculations with India ICS data.

Figure 5.1: IC percentage contributions to average log-employment. Whole Retail Sector



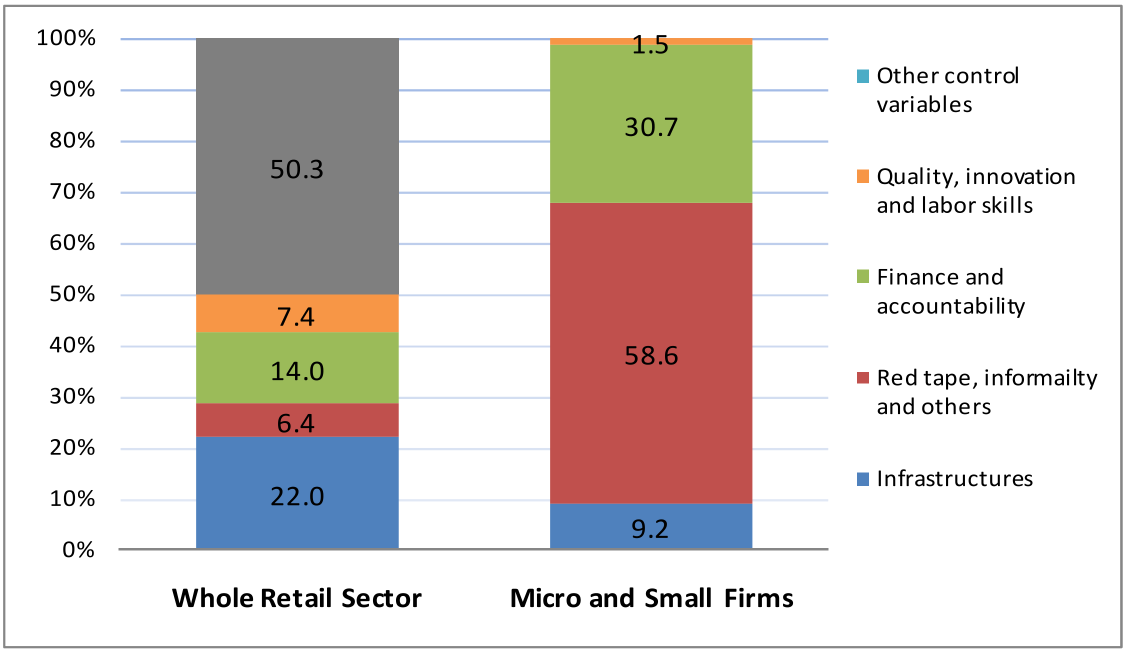
Note: Contributions computed according to section 3.
Source: Author's calculations with India ICS data.

Figure 5.2: IC percentage contributions to average log-employment. Micro and Small Firms.



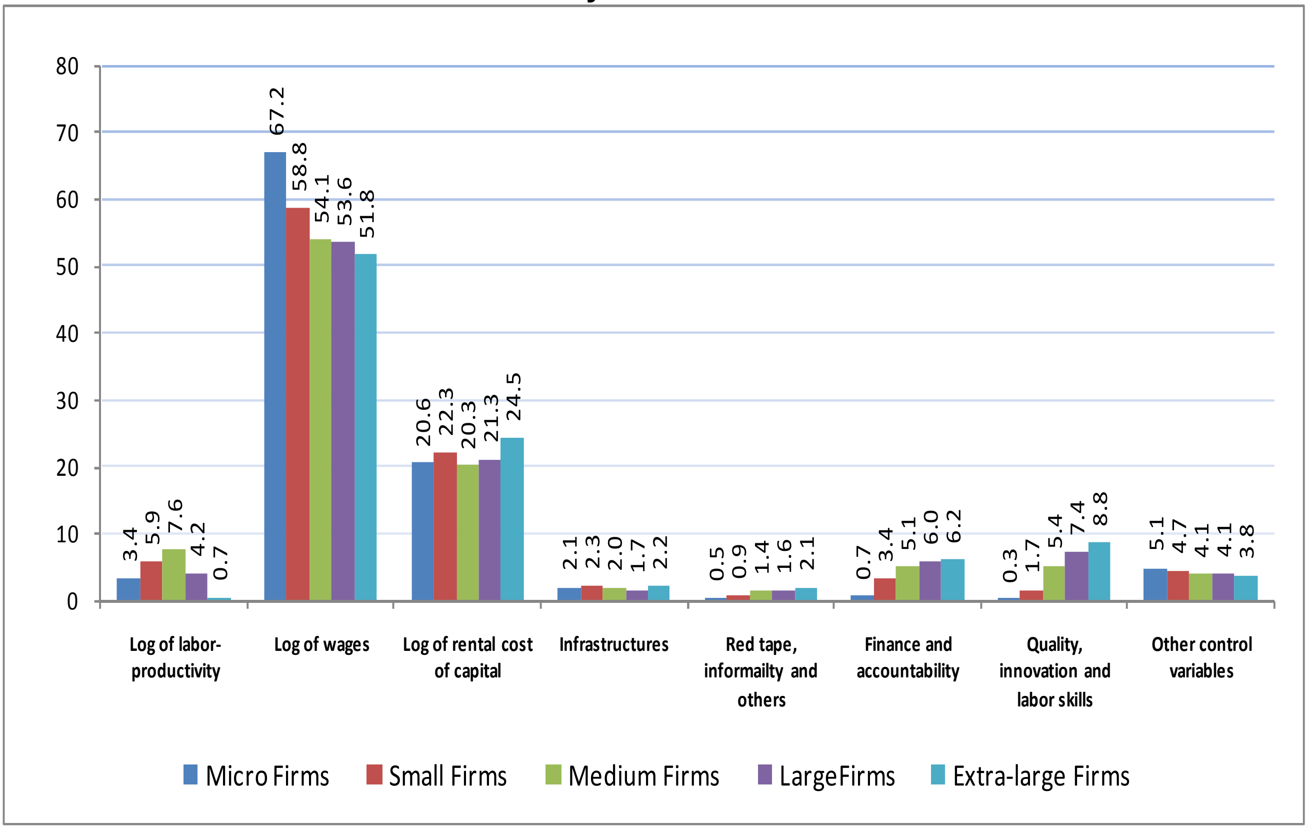
Note: Contributions computed according to section 3.
Source: Author's calculations with India ICS data.

Figure 5.3: IC Percentage Absolute Contribution on Average Employment Demand



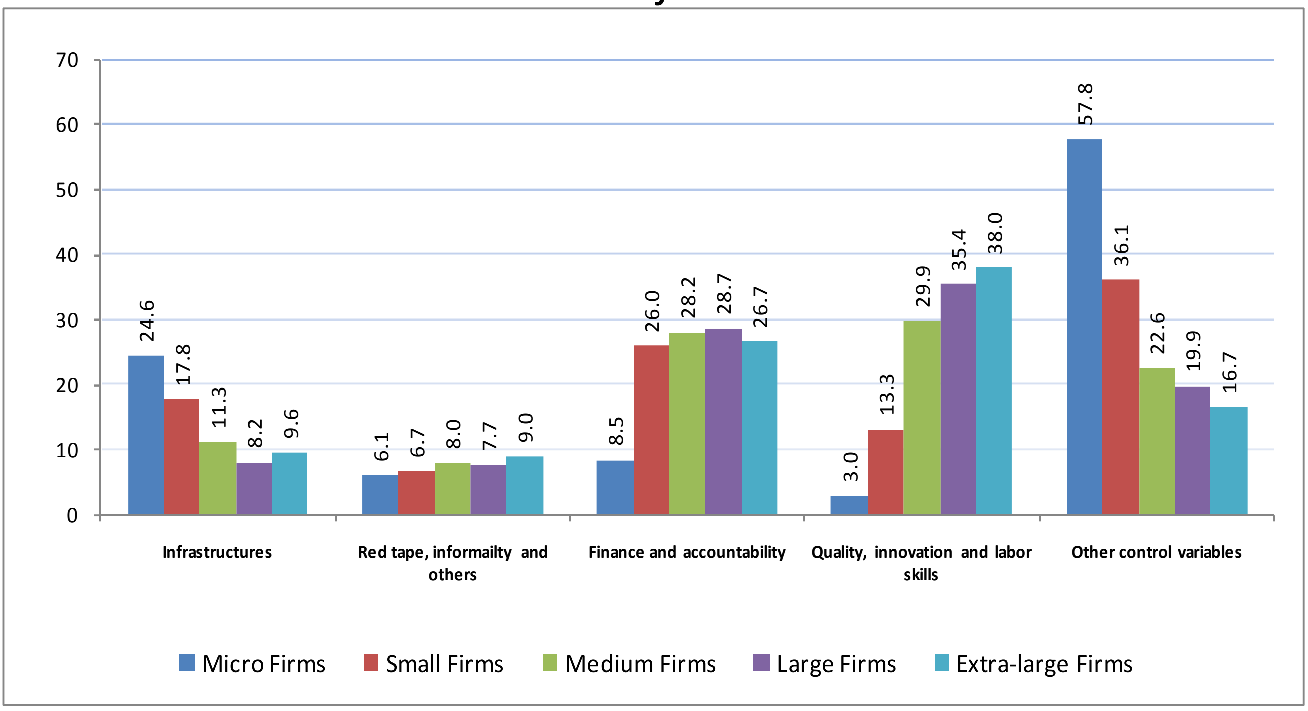
Note: Contributions computed according to section 3.
Source: Author's calculations with India ICS data.

Figure 5.4: Percentage Absolute Contribution on Average Employment Demand by Size



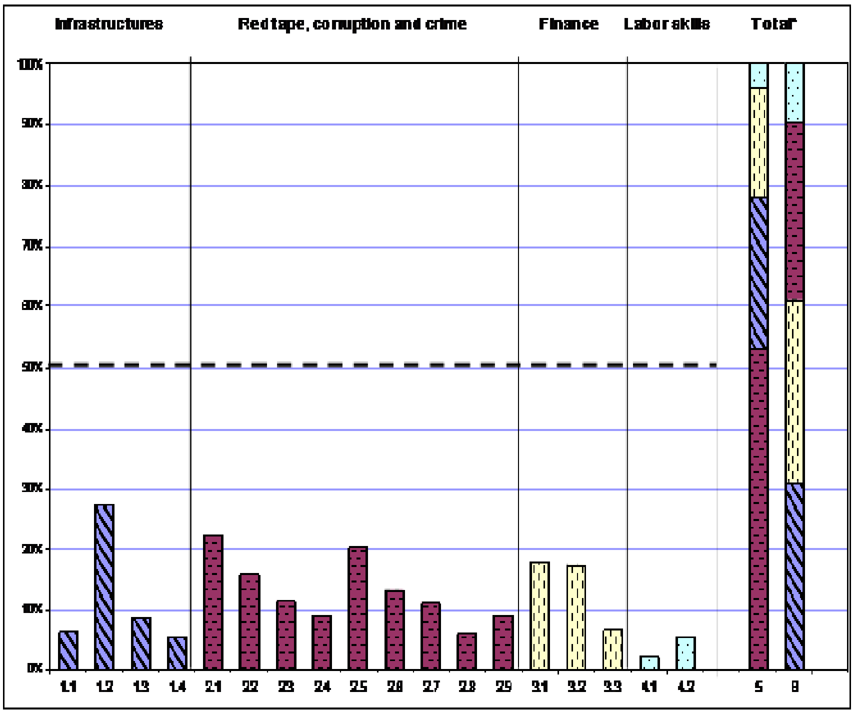
Note: Contributions computed according to section 3.
Source: Author's calculations with India ICS data

Figure 5.5: IC Percentage Absolute Contribution on Average Employment Demand by Size



Note: Contributions computed according to section 3.
Source: Author's calculations with India ICS data.

Figure 6.1: Firm's perceptions; percentage of firms that considers each one of the following problems as a severe obstacle to firms' economic performance.



1. Infrastructures.
1.1 Telecommunications
1.2 Electricity
1.3 Transportation
1.4 Customs and trade regulations

2 . Red tape, informality and others.
2.1 Corruption
2.2 Crime, theft and disorder
2.3 Anti-competitive or informal practices
2.4 Regulatory policy
2.5 Tax rates
2.6 Tax administration
2.7 Business licensing and operating permits
2.8 Regulation on hours
2.9 Regulation on pricing

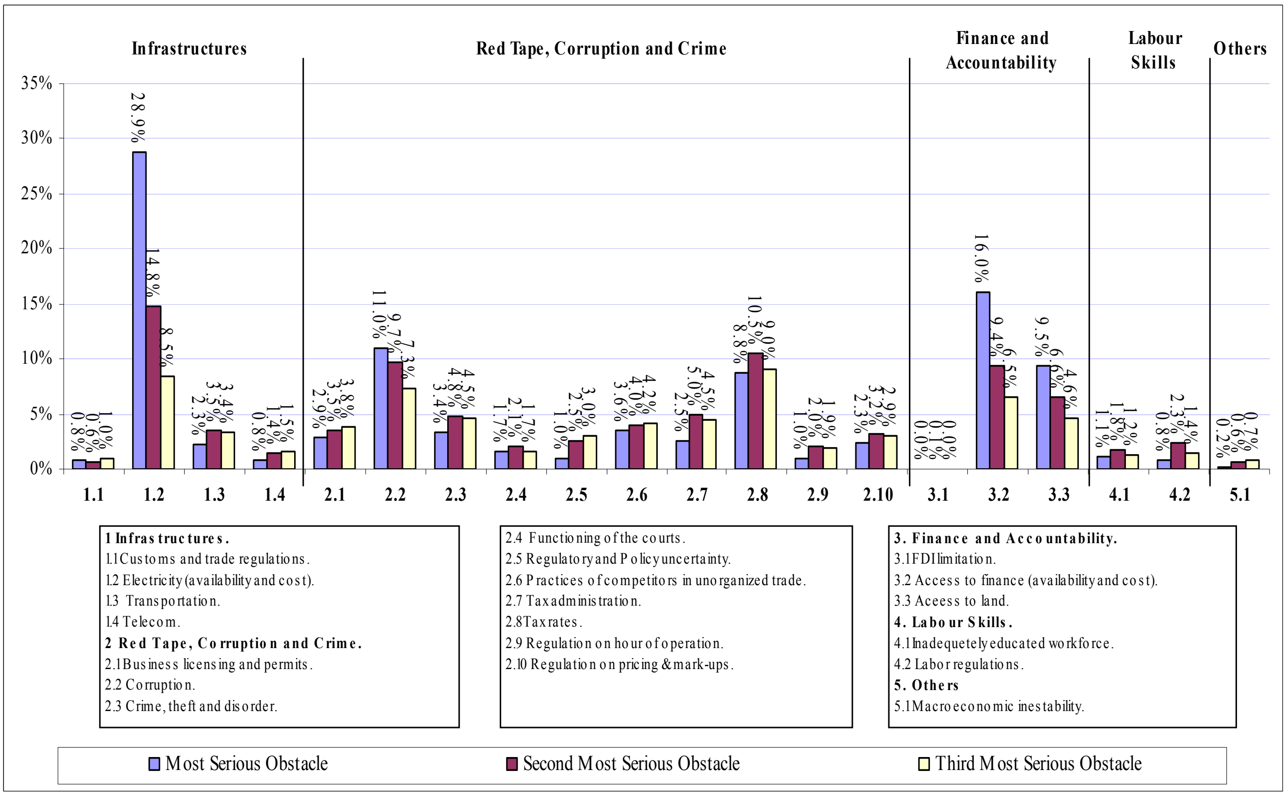
3. Finance.
3.1 Access to land
3.2 Cost of finance
3.3 Macroeconomic uncertainty

4. Labor skills.
4.1 Labor regulations
4.2 Skills and education of available workforce

5. Total relative weights.
6. Average group relative weights.
* (Totals are computed as the relative weigh of each group of perceptions over the sum of all perceptions' weights)

Source: Authors' calculations with India ICS data

Figure 6.2: Firm's perceptions; first, second and third most serious IC obstacles to firm's economic performance.



Source: Authors' calculations with India ICS data