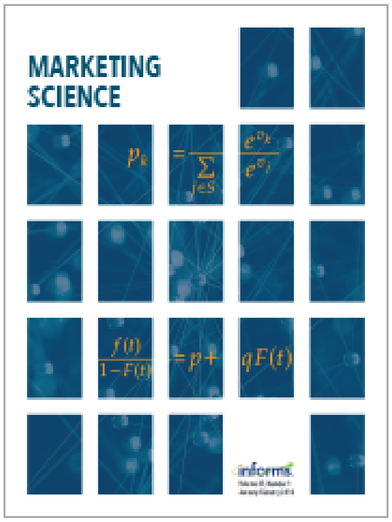


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The "Peter Pan Syndrome" in Emerging Markets: The Productivity-Transparency Trade-off in IT Adoption

K. Sudhir, Debabrata Talukdar

MARKETING SCIENCE
Vol. 34, No. 4, July–August 2015, pp. 500–521
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2015.0921

The "Peter Pan Syndrome" in Emerging Markets: The Productivity-Transparency Trade-off in IT Adoption

K. Sudhir
Yale School of Management, New Haven, Connecticut 06520, k.sudhir@yale.edu

Debabrata Talukdar
School of Management, State University of New York at Buffalo, Buffalo, New York 14260, dtalukda@buffalo.edu

Firms invest in technology to increase productivity. Yet in emerging markets, where a culture of informality is widespread, information technology (IT) investments leading to greater transparency can impose a cost through higher taxes and the need for regulatory compliance. The tendency of firms to avoid productivity-enhancing technologies and remain small to avoid transparency has been dubbed the "Peter Pan Syndrome." We examine whether firms make the trade-off between productivity and transparency by examining IT adoption in the Indian retail sector. We find that computer technology adoption is lower when firms are motivated to avoid transparency. Specifically, technology adoption is lower when there is greater corruption, but higher when there is better enforcement and auditing. So, firms have a higher productivity gain threshold to adopt computers in corrupt business environments that suffer from patchy and variable enforcement of the tax laws. Not accounting for this motivation to hide from the formal sector *underestimates* productivity gains from computer adoption. Thus, in addition to their direct effects on the economy, enforcement, auditing, and corruption can have indirect effects through their negative impact on adoption of productivity-enhancing technologies that also increase operational transparency.

Keywords: retailing; IT adoption; productivity; informal economy; emerging markets
History: Received: January 11, 2013; accepted: January 28, 2015. Preyas Desai served as the editor-in-chief and Gerard Tellis served as associate editor for this article. Published online in *Articles in Advance* May 29, 2015.

1. Introduction

For many businesses in emerging markets, information technology (IT) is a double-edged sword. On one hand, IT systems can help improve productivity and thus help firms gain a competitive advantage. Yet the same systems that improve productivity also increase transparency of transactions by leaving a clear audit trail. Such increased transparency makes it easier for the government to collect taxes and enforce regulatory compliance by bringing these transactions into the formal sector of the market, thus potentially increasing the firm's cost of operations, relative to those who do not use IT systems. In emerging markets, where enforcement is patchy and corruption is rampant, firms who keep much of their transactions in the informal sector can therefore gain a competitive advantage.[1] In such settings, the gains in productivity from adoption of IT are moderated by the attendant costs of making the transactions subject to taxation and regulatory compliance (Bird and Zolt 2008, Johnson et al. 2000, Mishra et al. 2008, Sinha 2003). At the margin, firms may therefore limit investments in IT, to the detriment of overall productivity, especially if their beliefs about the relative magnitude of productivity gains versus transparency costs are underestimated (Gatti and Honoratti 2008, Smith 2013). This tendency of firms in emerging markets to shun growth and remain small at the expense of efficiency, technology adoption, and innovativeness to avoid taxes and regulatory scrutiny has been dubbed the "Peter Pan Syndrome."[2] Sunder (2012) summarizes the dilemma in the context of domestic Indian retailers' reluctance to modernize through

[1] The informal or grey economy is here defined as trade, services or production, i.e., noncompliant in any aspect(s) of company registration, tax declaration/payment, business regulation (e.g., employer's national insurance, public/employer's liability insurance), and/or licensing requirements for the specific trade (e.g., health and safety certificate).

[2] In an article titled "The Peter Pan Syndrome," The Economist (May 17, 2014) states: "Manuel Milano of the Mexican Competitiveness Institute, a think-tank, calls this a 'Peter Pan System' in which firms prefer to stay small than to grow, mostly because of tax and regulation. It is easier to fly under the radar when you are microscopic." The article goes on to discuss the large opportunity costs of firms for remaining small. These costs include higher interest rates from banks, and major reductions in efficiency, technology, and innovation.

IT systems: "The system that serves to manage large retail organizations is also convenient for tax payment and collection... Indian retailers can and should break out of the self-defeating confines of the beliefs about the profitability of tax evasion."

There is some intuitive appeal to the conjecture that transparency concerns might impede IT adoption among emerging market retailers; however the conjecture has not received empirical scrutiny. Just as important, the productivity-enhancing benefits of IT adoption in emerging markets cannot be taken for granted. For example, it is possible that, given the low cost of labor and the lack of complementary infrastructures, the gains through productivity enhancement from IT adoption by retailers in emerging markets may not be sufficiently high to warrant such adoption. For IT to help improve productivity, the business ecosystem and organization should be able to take advantage of the technology. For example, in the absence of supply chain and cold chain infrastructure in emerging markets, the value of computers for efficient supply chain management may be quite limited. Similarly, when a retailer's employees are older, untrained, and unfamiliar with using IT systems, installing those systems will not lead to productivity gains. This is particularly relevant. Even within an advanced high income economy such as the United States, there was much academic debate until the mid-1990s as to whether IT in fact improves productivity.

For instance, much of the early research on IT productivity claimed a IT-productivity paradox in that it was not possible to reject the hypothesis that computers add nothing to total output (e.g., Loveman 1994), or found that the marginal costs exceeded marginal benefits (Morrison and Berndt 1990).[3] It was not until Brynjolfsson and Hitt (1996) showed through detailed firm-level survey data that, dollar for dollar, spending on computer capital created more value than spending on other types of capital, that the tide began to turn and researchers demonstrated that IT does increase productivity. The literature discusses two reasons for the divergence of results. First, the results reporting insignificant effects were from the 1970s when IT productivity may have been lower. Second, as discussed earlier in the examples, complementary infrastructure and the organizational redesign necessary to exploit IT may not have been present (Commander et al. 2011). Because complementary infrastructure may be inadequate and firms could still be in the low productivity part of the experience curve in emerging markets, the conjecture that IT improves productivity in this market deserves systematic empirical scrutiny.

[3] Robert Solow, the Nobel Prize-winning economist, characterized the IT productivity paradox thus: "We see computers everywhere except the productivity statistics."

Our goal here is to empirically answer three questions about the use of IT by businesses in emerging markets: First, do operational transparency concerns impede IT adoption by businesses? Second, does IT adoption have a positive impact on productivity and how much? Of particular significance, the magnitude of the impact may be underestimated at the margin if firms with potentially high productivity gains do not adopt computers due to transparency concerns that impede such adoption. Third, how does this trade-off vary by the size of the firm? Do transparency concerns reduce IT adoption among smaller or larger firms? Do larger firms gain more in productivity than smaller firms through such adoption? We answer these questions using detailed firm-level survey data on 1,948 retail firms covering a broad cross-section of Indian states and cities. We augment this firm-level survey data with state-level data in terms of a number of relevant variables such as corruption level, minimum wage rates, and overall socioeconomic development indices.

India presents an ideal setting for the study of these questions. First, the retail sector is at an early stage of modernization. Labor is still relatively cheap. Complementary infrastructures are still not fully available. Hence, the productivity gains from IT adoption is, a priori, ambiguous, requiring systematic empirical analysis. Specifically, the minimum wage rates and literacy levels vary across states. This gives us state-level variation on the labor saving productivity benefits of using computers. Second, with high levels of corruption in India, the transparency concerns are especially acute. (India scores a poor 36 (out of 100) in the Transparency International (2012) report and is ranked 94th out of 176 countries.) Furthermore, given India's federal system of government where states have significant power, there is a considerable variation in the levels of corruption, enforcement, and auditing across different states. These variations are valuable in identifying the empirical link between IT adoption and transparency motivation. We conduct a falsification test to determine whether the link between IT adoption and transparency levels across states is not merely due to another unobserved factor that varies across states, but is correlated with transparency and technology adoption. Specifically, we test the link between generator adoption and transparency variables, as transparency concerns should not affect generator adoption. Consistent with our hypotheses, we find that (unlike IT adoption) generator adoption is not linked to transparency-related factors.

Furthermore, in evaluating the effect of IT adoption on productivity, there are obvious selection concerns because business computer adoption is not random. To assess these concerns, we use two approaches. First, we use propensity score matching (PSM) to ensure that inferences of productivity differences between adopters and nonadopters are between firms that are comparable in their propensity to adopt. We also test for potential selection on unobservables using a Rosenbaum bounds approach to determine whether unobservable factors related to computer adoption might drive the positive estimates of productivity effects. Second, we estimate a model of self-selection using transparency variables as instruments. Variations in corruption and enforcement levels across states and firms serve as exclusion restrictions in that they impact computer adoption by firms, but do not directly impact firm revenues.

Our key findings are as follows: (1) At the margin, higher corruption levels are related to lower computer adoption; (2) Better regulation enforcement increases computer adoption because it creates a level playing field across firms, thus reducing transparency concerns; (3) Generators increase productivity, but as one would expect, their adoption is not affected by transparency concerns; (4) Computer adoption increases store productivity on average by about 50% to 70%. (The effects of transparency on computer adoption and the impact of computer adoption on productivity are both greater for larger firms than for smaller.); and (5) Not taking into account the endogenous effects of transparency related variables on computer adoption underestimates the productivity gains from IT adoption. This suggests that productivity estimates in emerging markets with nontransparent environments should account for such concerns.

Our results have obvious implications for policy makers. Our results show that corruption and lax enforcement of tax laws lead to direct losses in tax revenues. Moreover, indirect losses due to productivity decrease from reduced adoption of productivity-enhancing systems that increase transparency. From a marketing perspective, our results show that transparency concerns will reduce the market size of productivity-enhancing products (e.g., computers, cash registers (which maintain records in memory), and credit card machines) that also increase transparency. Furthermore, they suggest that marketers should use variables measuring corruption, enforcement levels, and audit mechanisms as predictors for market potentials for such products.

The rest of the paper is organized as follows. Section 2 provides background on the Indian retail sector, informal sector issues in emerging markets, and literature on IT productivity. Section 3 describes the data. Section 4 describes our empirical analysis and the results. Section 5 summarizes this paper, its goals and limitations, and areas for future research.

2. Background

We position this paper against two streams of literature: the IT-productivity relationship and the culture of informality in emerging markets. Finally, we discuss why the Indian retail sector is a particularly appropriate setting to study productivity-transparency trade-off.

2.1. IT Adoption and Productivity

As discussed in §1, the link between IT adoption and productivity was the subject of much controversy in the 1980s and early 1990s. Early analysis using firm-level data from 1978–1982 found no evidence of productivity increases (Loveman 1994, Barua et al. 1995). It is possible that productivity gains were small in the early stages of IT adoption. Others have argued that the inability to detect productivity gains could stem from aggregation and measurement bias (Brynjolfsson and Hitt 1996, Stiroh 2010). This productivity paradox was resolved through analysis of later data between 1987 and 1992 by Brynjolfsson and Hitt (1996). Since then a number of studies have found a strong and positive association between IT adoption and productivity (e.g., Ichniowski et al. 1997, Black and Lynch 2001, Bartel et al. 2005). At the same time, the magnitude of IT productivity gains varies significantly across countries, with estimates for European economies far lower than for the United States (Basu et al. 2003, Jorgenson 2001, Stiroh 2002). Clearly, one needs complementary logistics and a supportive regulatory environment for the effective use of IT in a national economy (Commander et al. 2011). Emerging markets may lack these factors thus potentially limiting productivity gain from IT. Often organizations must be redesigned to support IT. Because this redesign lags IT adoption, the benefits of such adoption may not be immediately apparent.

Note also that the link between firm size and productivity gains from IT is theoretically unclear. Whereas larger firms have more complex coordination needs that may facilitate greater productivity gains (Dasgupta et al. 1999), smaller firms may be more flexible to take better advantage of IT (Morgan et al. 2006). Not surprisingly, empirical results remain mixed. Most papers report a positive relationship (e.g., DeLone 1981, Fabiani et al. 2005, Morgan et al. 2006, Thong 1999), whereas some report insignificant (e.g., Lefebvre et al. 2005, Love et al. 2005) and negative relationships (e.g., Dewett and Jones 2001, Harris and Katz 1991).

2.2. Culture of Informality

A culture of informality, where firms keep business output hidden or opaque from the formal system of monitoring and thus avoid government taxation and regulation, varies across economies (Dabla-Norris et al. 2008). The share of informal business activities is estimated at between 10% and 20% of GNP for developed countries; it ranges from 33% to 50% for developing countries (Schneider and Enste 2002). Note that this practice is not limited to firms in the informal sector, especially in emerging markets. A report by McKinsey Global Institute (Farrell 2004) notes: "The informal economy is not just the unregistered street vendors and tiny businesses that form the backbone of marketplaces in Asia and other emerging markets. It includes many established companies, often employing hundreds of people, in industries as diverse as retail, construction, consumer electronics, software, pharmaceuticals and even steel production."

Firms prefer informality as it helps them avoid taxes and costly regulation. Unilateral avoidance of taxes becomes a competitive advantage when firms are unlikely to be caught and punished. For example, when corruption is high or enforcement is patchy, tax avoidance is feasible through paying bribes. Furthermore, by keeping tax-related operational activities informal and avoiding transparency-enhancing technologies, firms can reduce the level of electronic trail government officials can have in demanding bribes[4] (Mishra et al. 2008, Russell 2010). Unfortunately, the culture of informality leads to a vicious cycle of further tax avoidance and increased informality. Governments are forced to increase tax rates from the smaller (compliant) firms, which moves those firms toward noncompliance (e.g., Azuma and Grossman 2008, Dabla-Norris et al. 2008, Marcouiller and Young 1995). By contrast, when the enforcement environment is excellent and there are auditing mechanisms to make tax avoidance more difficult, firms are less likely to be in the informal sector and less motivated to avoid transparency-enhancing technologies. This is because computers provide productivity-enhancing benefits, but do not put the firm at a competitive disadvantage, and better enforcement ensures a level playing field for all players.

There is also face validity that business computerization increases operational transparency and facilitates better enforcement by creating easily detectable digital traces of taxable business activities through a transparent recordkeeping system (Friedman et al. 2000, International Tax Compact 2010, Russell 2010). For example, the governments of Bangladesh (The Daily Star 2007), China (People's Daily 2000), and Ethiopia (Mesfin 2012) recently mandated use of computerized systems to facilitate easy enforcement and minimize tax evasion.[5]

2.3. Choice of Setting: Indian Retail Sector

The Indian retail sector is the fifth largest in the world with a current market size of about US$500 billion and average growth rates of between 8%–10%. Yet the Indian retail sector lags behind peer emerging markets such as China in the adoption of modern management technologies and IT systems to facilitate retail business practices (Reardon and Gulati 2008, Sunder 2012). Thus, the retail sector in India is an ideal setting for studying the productivity-transparency tradeoff. Many argue that the sector is well positioned for productivity gains from IT adoption through improvement in inventory management, pricing, and customer relationship management (Foster et al. 2002, Sunder 2012). Yet the low rate of IT adoption could be because the Indian environment is not conducive to productivity gains from such adoption. For example, lack of complementary infrastructure (e.g., logistics and supply chain, road infrastructure, etc.) may limit the productivity gain from IT adoption; low labor costs may also limit such potential gains.

Transparency issues are a reality in India, where there is an endemic national culture of corruption. Transparency International (TI) found that more than half of those surveyed had first-hand experience paying bribes or peddling influence to obtain a public office job. India was ranked 94th among 176 countries for lack of transparency (TI 2012). India tops the worldwide list for "black money" with almost $1.456 billion secreted in Swiss banks (Nayar 2011, Rao 2010). This amount is 13 times the country's total external debt. The popular press is replete with articles noting that "tax evasion is a national sport" for businesses and individuals (Chopra 2011, Dhara and Thomas 2011). Understanding the relative importance of productivity and transparency as it relates to the low rate of computer adoption can be a critical aid to policy prescriptions for improving productivity in one of the world's largest retail markets.

3. Data

We collated the data necessary for the analysis from multiple sources. We first discuss the sources and then provide descriptive statistics of the variables used in our empirical analyses.

[4] Miller and Tucker (2014) find hospitals in the United States are one-third less likely to adopt electronic medical records (EMR) systems in those states that allow search and use of electronic records in litigation cases, even though EMR systems enhance operational productivity and cost efficiency. Their results suggest that, even in developed countries, transparency concerns can reduce IT adoption.

[5] People's Daily (2000) reported: "China has stepped up its efforts to fight against tax evasion by requiring selected companies to print invoices using a computerized system connected to taxation authorities."

3.1. Data Sources

Our primary data source for this study is a large scale World Bank survey of Indian retailers conducted in 2006. As part of its private sector development project and research initiative, the World Bank conducts regular surveys of individual firms in many developing countries. Although these surveys are primarily used to guide internal bank policies, they have also been used to address academic research questions in economics and finance (e.g., Angelini and Generale 2008, Cull and Xu 2005). Amin (2010) uses data from the 2006 survey to study the effect of labor regulation on computer adoption.

The survey consists of a stratified random sample of 1,948 retail stores operating in the formal sector in 16 major states and federal territories across 41 Indian cities. The National Industrial Classification (NIC) groups Indian retailers into those operating through registered stores and those who usually operate informally from home (NIC 1998, Industry Division 52). All stores in our sample belong to the former group.

The sampling was carried out with a first level stratification of three segments by retail store type: (i) traditional stores, including general and department stores, grocers, drugstores, food stores, etc.; (ii) consumer durable stores, which carry televisions, home appliances, etc.; (iii) modern format large stores that are part of a shopping complex. These three store types account for 64%, 26%, and 10% of the sample, respectively. Within each store type segment, a secondary stratification was based on operation size. The overall sample size was designed to minimize the standard error in the sample variables, given the available resources for each surveying stratum.

The survey was conducted by the Indian unit of a well-respected international market research firm. Personal interviews were conducted with store managers who were told that the goal of the survey was to gather opinions about the Indian investment climate for the retail sector. They were also told that the information obtained would be strictly confidential; neither their names nor the names of their businesses would be used in any document based on the survey. Data was collected on a variety of store characteristics including annual sales, key operational costs, employment, availability of infrastructure, access to finance, etc. The data also reports on the store manager's perceptions about various aspects of the business climate including competition and the culture of corruption.

We augment this store-level survey data from the World Bank with relevant state-level data from other sources. Specifically, the state-level corruption index is obtained from the TI Indian Corruption Study, which was released in October 2005. This is one of the largest corruption surveys ever conducted, with a total of 14,405 respondents from 151 cities and 306 rural areas in 20 Indian states. We also collected data from Indian government sources on three other state-level variables to capture state-level differences that can affect IT adoption. One of these variables is labor cost. Computers and electronic cash registers can replace (1) competent and experienced accounting and stock-keeping staff who use traditional manual accounting books, and (2) experienced and trusted cashiers who are proficient in mental computations to total up bills and provide change. Hence, higher labor costs make automation through computers to increase productivity more appealing (Amin 2010). We operationalize labor costs through the minimum wage rates in the retail services sector across states. These rates are set under the Shops and Establishments Act (SEA) of India. We use data from the Indian Labor Bureau 2001 report, which is the closest year to our 2006 World Bank survey for which data was available for all of the states in our sample. Though this 2001 report data differs from actual 2006 wages, we believe the relative values will be comparable.

Another state-level variable on which we collect data is the adult literacy rate as a proxy for relative education level differences across states. In states with a less educated workforce, it would be harder to find employees who can use computers effectively; this would lead to lower IT adoption. A less educated public may also tolerate more corruption. We use the average states' adult literacy rates from the Indian government 2001 and 2011 census data. This average is likely to be close to the 2006 literacy level. Finally, we use the Human Development Index (HDI) to capture differences in socioeconomic development across states; lower development can inhibit IT adoption and foster corruption. As these two state-level variables impact corruption and IT adoption, it is important to control for them in isolating the direct effect of corruption and other transparency metrics on adoption. We obtain the HDI data for our sample states from the 2011 India Human Development Report (Government of India 2011), which computes the index values based on 2007–2008 national survey data. The HDI for a state is a composite relative indicator of the socioeconomic development stage for the state along three key dimensions, i.e., education, health, and income levels of its population.

3.2. Descriptive Statistics

Table 1 shows summary statistics of the variables used in the empirical analyses.[6] We measure store performance by gross revenue generated in the latest fiscal year, normalized for size of the retailer in terms

[6] Additional background information, including inclusion rationale and operationalization, for some of the variables are in the online appendix (available as supplemental material at http://dx.doi.org/10.1287/mksc.2015.0921).

Table 1 Summary Statistics of the Analysis Variables[1]

Description	*N*	Mean	SD
Store level performance measures			
Gross revenue generated—latest financial year (Rs. in million)	1,918	1.900	4.081
Revenue net of operational costs—latest financial year (Rs. in million)	1,849	1.619	3.105
Productivity-enhancing technology adoptions			
Business computer (0 = No; 1 = Yes)	1,948	0.167	0.373
In-store electricity generator (0 = No; 1 = Yes)	1,948	0.296	0.456
Corruption factors discouraging transparency			
Self-assessment of % of revenue typically reported by peers for tax purposes	1,669	58.148	39.040
Self-assessment of % of revenue typically used to bribe regulatory agencies	1,808	0.835	2.366
Transparency International (TI) Corruption Index at the state level (1–10)[2]	1,948	4.811	0.769
Enforcement factors encouraging transparency			
Number of times the store was inspected last year by state regulatory agencies	1,948	1.512	3.580
Store has an external auditor (0 = No; 1 = Yes)	1,914	0.302	0.459
Perceived consistency in state's regulatory implementations (1 = Low; 6 = High)	1,948	3.096	0.608
Other state level variables			
Labor cost in terms of minimum wage rate (Rs.)[3]	1,948	73.383	13.421
Literacy rate (percentage)[4]	1,948	72.922	8.335
Human Development Index (0–1)[5]	1,948	0.509	0.095
Electricity power supply related factors			
Faced power outage over the last year (0 = No; 1 = Yes)	1,944	0.829	0.377
State's power supply as a perceived obstacle to business (0 = No; 4 = Severe)	1,948	1.635	0.468
Store size and age characteristics			
Floor area of the store (sq. ft.)	1,938	599.811	3,553.710
Number of full time employees at the store	1,948	5.722	24.557
Age of the store (years)	1,948	14.478	12.796
Store management and ownership characteristics			
Experience of the store manager (years)	1,948	12.948	9.803
Ownership concentration (% of store owned by the largest owner)	1,948	96.073	16.056
Government owned store (0 = No; 1 = Yes)	1,948	0.011	0.103
Store finance, in-store security, and competitive factors			
Business bank account (0 = No; 1 = Yes)	1,940	0.639	0.481
Overdraft facility (0 = No; 1 = Yes)	1,921	0.223	0.416
In-store security system (0 = No; 1 = Yes)	1,947	0.266	0.442
Perceived level of price competition (0 = Low; 1 = High)	1,901	0.376	0.484
Inventory level maintained for the main product (days)	1,948	11.582	16.167

[1]Unless specifically noted, the data source for a variable is the 2006 World Bank survey of Indian retailers.
[2]*Data source.* "*India Corruption Study* 2005," Transparency International, Centre for Media Studies, New Delhi, India.
[3]*Data source.* "*Report on the Working on the Minimum Wages Act of* 1948 *for the Year* 2001," Labor Bureau, Government of India. Accessed at http://www.labourbureau.nic.in/MW2K1%20Main%20Page.htm.
[4]*Data source.* Average of the states' adult literacy rates from the Indian Government 2001 and 2011 census data.
[5]*Data source.* "*Human Development Report* 2011," Government of India.

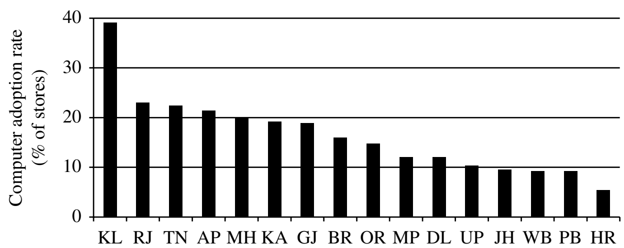
of employees and store area. Specifically, our performance metrics are gross annual revenue per employee (labor productivity) and gross annual revenue per square feet (floor area productivity).[7] For our sample stores, mean gross revenue is Rs. 1.90 million. Revenue net of operational costs is Rs. 1.62 million. The median gross revenue is Rs. 0.50 million. For firms at or below the median, the average revenue is Rs. 0.22 million, whereas for firms above the median, the average revenue is Rs. 3.79 million. For our measure of labor productivity, the mean and median values are Rs. 0.55 million/employee and Rs. 0.25 million/employee, respectively. For the floor area productivity measure, the mean and median values are Rs. 7,450/sq. ft. and Rs. 3,330/sq. ft., respectively.

We consider adoption of two productivity-enhancing technologies by each retail store in our sample. These technologies are a business computer and an in-store electricity generator. Although only 17% of the stores have business computer systems, 30% own a generator, while 27% own an in-store security system. This suggests that the absence of computer technology may not be due entirely to financial constraints. With 83% of the stores facing power outages in the previous year, perhaps greater generator adoption was

[7] In addition, because the World Bank survey collected data on some key annual operational costs, namely, labor, electricity, communication services, and rent or loan payment on land/building, equipment, and furniture, we also tested the robustness of our results for productivity measures based on gross revenues net of those costs for the latest fiscal year. The key results are qualitatively identical and are available in the online appendix.

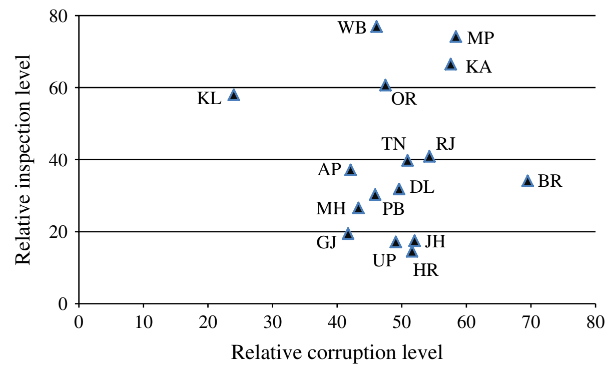
Figure 1 **Adoption Level of Business Computers By Store Type (% of Stores)**

Figure 2 **Adoption Level of Business Computers By State**



Note. AP, Andhra Pradesh; BR, Bihar; DL, Delhi; GJ, Gujarat; HR, Haryana; JH, Jharkand; KA, Karnataka; KL, Kerala; MP, Madhya Pradesh; MH, Maharashtra; OR, Orissa; PB, Punjab; RJ, Rajasthan; TN, Tamilnadu; UP, Uttar Pradesh; WB, West Bengal.

Figure 3 **(Color online) Relative Inspection versus Corruption Levels Across the Sample States**



optimal. This also reflects the tremendous loss of efficiency and wasted capital in emerging markets, where the absence of infrastructure (i.e., power) necessitates what might be an otherwise wasteful investment in in-store generators.

We consider both state- and firm-level perceptions of corruption as it is an experiential phenomenon that occurs away from public glare. Hence, even though perceptions of corruption by individual firms within a state will be correlated to the aggregate state-level index, it will also vary across firms because of differential experiences in the context of individual business operations. For example, officials in local regulatory offices who deal with a particular firm are likely to be different in their propensity toward corruption. Similarly, the peer group of firms, whose actual or perceived operational practices shape firm perception of corruption prevalence, will differ across firms, even within the same geography.

The data show that stores operate in business environments that vary in terms of corruption-related factors expected to discourage operational transparency, as well as in terms of regulatory enforcement related factors expected to encourage transparency. For the states included in the World Bank survey, the values of the TI corruption index (measured on a 1–10 scale) range from a low of 2.40 (Kerala) to a high of 6.95 (Bihar). Other variables also vary significantly across sample states. For example, the minimum wage rate varies from a low of Rs. 42.50 to a high of Rs. 99.70, with an average value of Rs. 72.38. Figures 1 and 2 show the sample distribution of adoption level of computers by store type and state, respectively. Whereas the overall adoption level is low at 16.8%, there is a significant variation across both store type and state. Figure 3 shows our primary explanatory variables for why computer adoption varies by (1) enforcement, and (2) corruption across the 16 states in our sample. There is a substantial variation across states in terms of enforcement and corruption.

In terms of store-specific characteristics, the average number of employees in a store is approximately six, but there is a substantial standard deviation around the mean. The median number of employees is two. For firms at or below the median number, the average number of employees is 1.3. For firms above the median number of employees, the average number of employees is 12. The average store size is about 600 sq. ft., but here also there is a large standard deviation around the mean. The median size is 150 sq. ft. For firms at or below the median, the average size is 90 sq. ft. For firms above the median, the average size is 1,167 sq. ft. The average age of the store is 12 years. Here the standard deviation is less than the mean, and the mean and median are roughly the same, unlike for size and number of employees.

The average experience of the store manager is 13 years. Most stores are single proprietor. The mean value of the store shares held by the largest shareholder is about 96%. As expected, only 1% of stores are government-owned. Indian retailing is almost entirely a private sector activity, except for fair price shops that distribute staple groceries to the poor. Stores vary in their level of access to formal financing. About 36% of stores do not have bank accounts; 78% do not have access to overdraft facilities, which suggests significant credit constraints. Stores keep about 12 days of inventory for their main products.

4. Empirical Analyses

We begin with bivariate descriptive analyses to obtain a graphic sense of the nature of relationships among

Figure 4 (Color online) Corruption Level and Computer Adoption By State

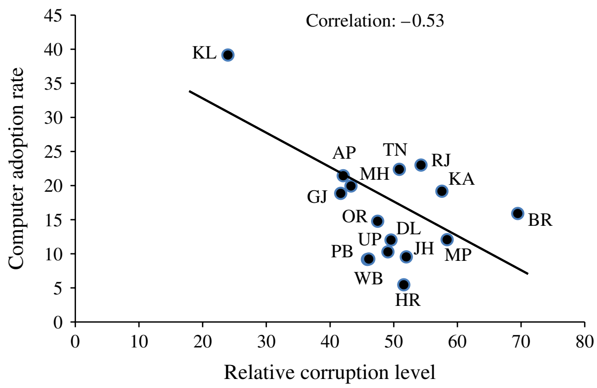
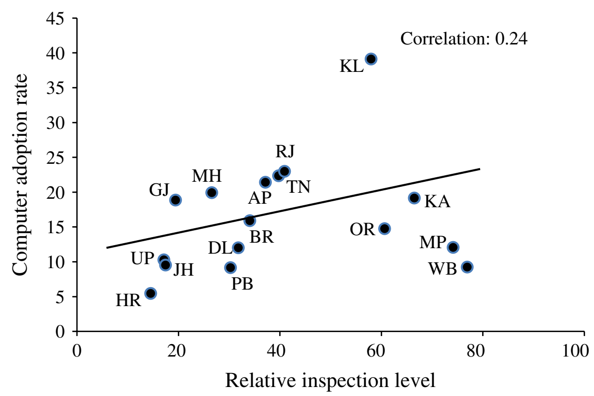


Figure 5 (Color online) Inspection Level and Computer Adoption By State



the key variables of interest. Next we present relevant statistical analyses that control for other variables that can impact the outcome variables. We also account for potential endogeneity concerns in estimating the productivity impact of computer adoption, which is an endogenous variable.

4.1. Bivariate Relationships

We first report the relationship between computer adoption levels and transparency and enforcement. Figures 4 and 5 show the state-level scatter plots of computer adoption levels with corruption and regulatory inspection levels. The correlation between computer adoption and corruption levels is −0.53. This is consistent with the premise that the higher the overall culture of corruption, the higher the propensity towards business tax evasion and thus the lower the incentive to adopt transparency-enhancing business computer technology. Similarly, the correlation of 0.24 between adoption and inspection level is consistent with the premise that the higher the regulatory enforcement, the lower the propensity towards business tax evasion and the higher the incentive to adopt transparency-enhancing computer technology.[8]

Figure 6 Computer Adoption by Median Split of States By Corruption Level

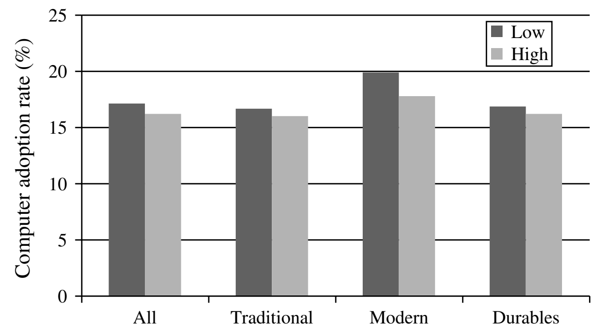
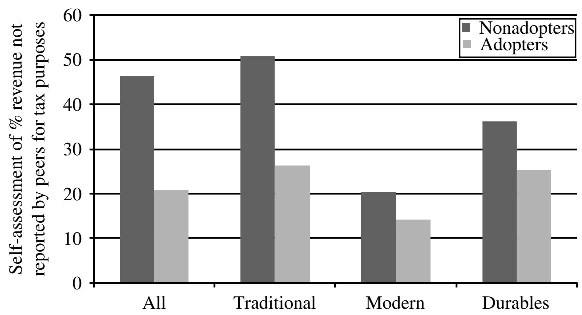


Figure 7 Computer Adoption by Median Split of States By Inspection Level

We further test whether the overall correlations reported above hold within more specific subgroups. First, we do a median split of states by the level of corruption to determine whether computer adoption is lower in states with higher corruption levels. We also determine whether the relationship holds within subgroups of retailers (traditional, modern, and durables). The results are reported in Figure 6. We find that the computer adoption rate is lower by 5% to 12% in higher corruption states. Interestingly, the gap is larger among modern stores, which suggests that modern retailers are most strategic about transparency concerns when adopting computers. Figure 7 shows a similar graph for computer adoption rates as a function of enforcement levels. The gap is larger here. States with better enforcement have 30% to 40% higher adoption rates. Clearly, one has to control for other variables to quantify the effect of transparency and enforcement on computer adoption. Still, there is prima facie evidence that higher enforcement and

[8] Because Kerala (KL) and Bihar (BR) appear to be outliers in Figure 4, we assessed the robustness of our results by dropping Kerala and Bihar from the states included in the analysis. All of the reported relationships in the paper continue to hold. The key results without the outlier observations are available in the online appendix.

Figure 8 **Perceived Dishonesty Among Peers—By Computer Adoption Across Store Types**



better governance are correlated with lower transparency concerns.

To further assess the transparency concern, we used another variable that indicates the competitive disadvantage of tax evasion. This variable is store managers' perceptions of the level of dishonesty among their peers in terms of hiding revenues for tax evasion. Figure 8 reports how computer adoption varies among different store types, based on the managers' perception of perceived dishonesty among peers. For all store types, managers who have not adopted computers in their stores believe that there is a higher level of tax evasion in the industry.

Finally, Figure 9 reports the relationship between productivity and computer adoption using our two productivity metrics: the revenues per employee (labor productivity) and revenues per square feet (floor area productivity). The graphs show that productivity is higher for stores that adopt computers. Obviously, other variables need to be controlled for and there are potential selection concerns. We address these in the subsequent statistical analyses.

4.2. Computer Adoption

We begin by discussing the findings from the probit computer adoption regression shown in Table 2. The first set of results is for the full sample. The first column excludes gross annual revenue, whereas the second column includes it. The results in both columns are consistent with our hypotheses about transparency variables, i.e., corruption, enforcement/audits, and regulatory consistency. The corruption variables are negatively related to adoption, whereas enforcement/audits and regulatory consistency are positively related to adoption. Thus our primary hypothesis, i.e., that computer adoption is systematically correlated with transparency concerns is supported.

As expected, higher labor costs are positively related to computer adoption as firms substitute technology for labor. Computer adoption is positively related to generator adoption, suggesting a positive correlation in preferences for productivity-enhancing technologies. Furthermore, as expected, from a cost affordability perspective, computer adoption is positively correlated with a store's gross annual revenue.[9] In terms of control variables such as state literacy rates (or HDI),[10] and other store and management characteristics, they generally have the right sign. Higher literacy rates (or HDI) are correlated with greater computer adoption. Larger stores are more likely to adopt computers. Interestingly, older stores and managers with greater experience are less likely to adopt computers. This suggests (not surprisingly) that experience is negatively related to new technology adoption. This negative relationship with experience potentially captures the relative discomfort of older managers with computers. In §4.6 we revisit the net effect of managerial experience on productivity based on our self-selection model (SSM) estimation results. Finally, we find that ownership characteristics (private or government ownership) are not related to computer adoption. Also, interestingly, the power supply factors do not have a significant relationship. This suggests that generators and power back-up equipment are being used, as appropriate, to address power supply problems.

We conclude this section by assessing the heterogeneity in the effect of transparency concerns by firm size. We use a median split of firms by size and estimate the probit model of computer adoption separately for large and small firms. The direction of the estimates is qualitatively identical across firm size for all variables. However, the differences in the magnitude of the effects on small and large firms show that corruption and enforcement concerns systematically have a greater impact on larger firm adoption decisions than on smaller firms.[11] Specifically, corruption

[9] As shown in §4.4, gross annual revenue has a reverse causal link to computer adoption through the productivity link. We do not account for this endogeneity here; our primary interest is the transparency variables. In §4.6, we estimate the gross annual revenue and computer adoption equations as a simultaneous equations model accounting for self-selection with appropriate exclusion restrictions, where we control for the endogeneity of the revenue and computer adoption variables. Our results in Tables 8(a)–8(c) remain robust.

[10] At 0.85, the correlation between literacy rate and HDI in our sample of 16 Indian states is high. This is not surprising as education (proxied by literacy rate) is one of the three dimensions for HDI. Literacy rate is also strongly correlated with per capita income and life expectancy, which are the proxies for the other two dimensions, namely, economic development and health. To avoid multicollinearity, we use literacy rate and HDI separately in the regression analysis. The results are qualitatively identical. To conserve space, we present only the results using the literacy rate in the paper.

[11] Rather than estimate the probit separately for smaller and larger firms, one could estimate a pooled regression with interaction terms between the relevant variables with a large or small firm dummy. This would have highlighted whether the differences are significant, but would not have provided a direct estimate of the effect on small and large firms. We estimated the pooled regression with interactions. All differences discussed here are significant.

Figure 9 Productivity of Stores By Computer Adoption Status

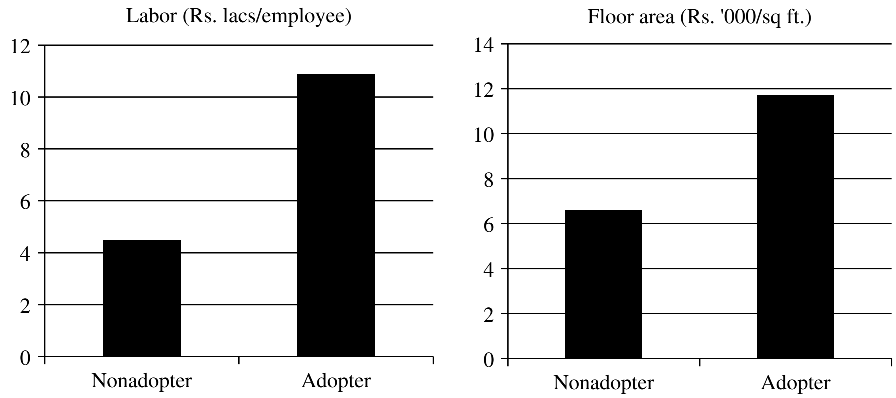


Table 2 Business Computer Adoption—Probit Regression

	All stores-I		All stores-II		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Corruption factors discouraging transparency								
State TI Corruption Index	−0.169**	0.076	−0.172**	0.083	−0.179**	0.085	−0.127**	0.061
Percent revenue spent on bribe	−0.020**	0.009	−0.026**	0.011	−0.026**	0.012	−0.017**	0.008
Perceived informality by peers	−0.003*	0.002	−0.003*	0.002	−0.002*	0.001	−0.003*	0.002
Enforcement factors encouraging transparency								
Regulatory inspections	0.033**	0.014	0.034**	0.015	0.038**	0.017	0.024**	0.010
External auditor	0.321**	0.139	0.293**	0.119	0.246**	0.106	0.448**	0.202
State's regulatory consistency	0.026**	0.012	0.028**	0.013	0.027**	0.012	0.026**	0.012
Labor cost and education level								
State minimum wage	0.041*	0.021	0.032*	0.017	0.038*	0.020	0.009*	0.005
State literacy rate	0.076**	0.036	0.071**	0.033	0.074**	0.035	0.062**	0.031
Electric power supply related factors								
Power outage	−0.058	0.091	−0.051	0.094	−0.055	0.101	−0.052	0.094
State power supply problem	−0.067	0.072	−0.076	0.078	−0.077	0.073	−0.070	0.072
Productivity-enhancing technology adoptions								
Generator	0.529**	0.254	0.508***	0.152	0.609***	0.152	0.541*	0.281
Store level performance measure								
Gross annual revenue	—	—	0.064***	0.022	0.062***	0.021	0.119**	0.056
Store characteristics								
Store size	0.312***	0.103	0.313***	0.098	0.225**	0.105	0.411***	0.142
Employee size	0.387***	0.088	0.401***	0.108	0.406***	0.106	0.293***	0.087
Store age	−0.185**	0.091	−0.186**	0.093	−0.222**	0.109	−0.182**	0.091
Store management and ownership characteristics								
Manager experience	−0.162**	0.078	−0.156**	0.073	−0.153**	0.071	−0.125**	0.058
Ownership concentration	−0.008	0.006	−0.005	0.005	−0.006	0.005	−0.007	0.008
Government owned	−0.274	0.469	−0.268	0.379	−0.269	0.385	−0.177	1.959
Fixed effects								
Store type and city	Yes		Yes		Yes		Yes	
Observations	1,501		1,501		734		767	
Model statistics	LL = −355.59		LL = −342.23		LL = −362.78		LL = −260.69	
	$\chi^2 = 441.36$ ($p = 0.00$)		$\chi^2 = 446.78$ ($p = 0.00$)		$\chi^2 = 467.43$ ($p = 0.00$)		$\chi^2 = 300.26$ ($p = 0.00$)	

Note. Standard errors are based on state level clustering.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

suppresses computer adoption more among larger firms, whereas enforcement increases computer adoption among larger firms. This suggests that the losses from corruption and the gains from enforcement are greater for larger firms. External auditors have a higher positive relationship with computer adoption among small firms relative to large firms.

4.3. Falsification Check: Generator Adoption

One interesting possibility is whether there are some unobserved characteristics that are correlated with transparency concerns and which might drive productivity-enhancing technology adoption. For example, states with high corruption might systematically engender nonadoption of productivity enhancing-technologies for reasons that might not be associated with transparency concerns. To address this, we conduct a falsification check. We choose a technology such as generators, whose adoption increases productivity, but that is not connected to transparency, and test whether its adoption is linked to transparency concerns. Furthermore, to assess face validity, we include literacy rate in the regression shown to impact computer adoption, but which should not directly impact generator adoption. The results of the generator adoption regression for the full sample, and for large and small firms are presented in Table 3.

The falsification check is validated. The transparency variables, i.e., corruption, enforcement, and regulatory consistency, are insignificant for generator adoption both in the aggregate as well as for small and large firms, separately. On the other hand, the factors related to electric power are shown to be highly significant. Larger stores are more likely to use generators. This could be because electricity infrastructure is worse in less developed areas. In that sense the electric power factors might be capturing some other transparency elements that we have not considered. However, state literacy rate is insignificant as literacy

Table 3 **Electricity Generator Adoption—Probit Regression**

	All stores		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE
Corruption factors discouraging transparency						
State TI Corruption Index	−0.080	0.094	−0.124	0.113	−0.044	0.218
Percentage revenue spent on bribe	−0.013	0.018	−0.031	0.028	−0.027	0.029
Perceived informality by peers	0.004	0.004	0.005	0.005	0.002	0.002
Enforcement factors encouraging transparency						
Regulatory inspections	0.026	0.031	0.009	0.022	0.075	0.083
External auditor	0.245	0.172	0.139	0.153	0.472	0.516
State's regulatory consistency	0.011	0.015	0.009	0.014	0.013	0.018
Labor cost and education level						
State minimum wage	−0.021	0.094	−0.026	0.113	−0.019	0.087
State literacy rate	0.043	0.034	0.045	0.039	0.040	0.035
Electric power supply related factors						
Power outage	1.568***	0.289	1.801***	0.344	1.417***	0.476
State power supply problem	0.728**	0.286	1.042**	0.433	0.586**	0.251
Productivity enhancing technology adoptions						
Computer	0.518***	0.141	0.662***	0.162	0.287***	0.110
Store level performance measure						
Gross annual revenue	0.068***	0.018	0.069***	0.022	0.127***	0.046
Store characteristics						
Store size	0.306***	0.070	0.275**	0.119	0.350**	0.164
Employee size	0.368***	0.075	0.441***	0.105	0.357***	0.133
Store age	−0.081	0.080	−0.004	0.108	−0.277	0.132
Store management and ownership characteristics						
Manager experience	0.266***	0.077	0.261***	0.103	0.339***	0.126
Ownership concentration	0.005*	0.003	0.004	0.003	0.006	0.007
Government owned	−0.254	0.465	−0.765	0.523	−0.378	0.313
Fixed effects						
Store type and city	Yes		Yes		Yes	
Observations	1,501		734		767	
Model statistics	LL = −512.63		LL = −294.97		LL = −230.14	
	$\chi^2 = 481.5$ ($p = 0.00$)		$\chi^2 = 368.2$ ($p = 0.00$)		$\chi^2 = 274.2$ ($p = 0.00$)	

Note. Standard errors are based on state level clustering.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4(a) Labor Productivity—OLS Regression

	All stores		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE
Productivity enhancing technology adoptions						
Computer	0.224**	0.097	0.253***	0.094	0.216**	0.093
Generator	0.244***	0.080	0.233**	0.097	0.267***	0.079
Store characteristics						
Store size	0.011**	0.005	0.012***	0.002	0.011***	0.003
Store age	−0.049	0.048	−0.085	0.070	−0.064	0.050
Store management and ownership characteristics						
Manager experience	0.113**	0.046	0.220***	0.065	0.098**	0.040
Ownership concentration	−0.003*	0.002	−0.002*	0.001	−0.006**	0.003
Government owned	0.600	0.461	0.735	0.561	0.236	0.738
Finance, in-store security, and competitive factors						
Bank account	0.316***	0.072	0.305**	0.124	0.319**	0.139
Overdraft facility	0.179**	0.080	0.189**	0.085	0.174**	0.077
In-store security	0.189**	0.076	0.209**	0.105	0.185**	0.085
Price competition level	0.095	0.065	0.125	0.097	0.036	0.083
Inventory level for main product	0.006***	0.002	0.007***	0.003	0.004***	0.002
State level educational factor						
Literacy rate	0.022**	0.011	0.019*	0.010	0.021**	0.009
Fixed effects						
Store type and city	Yes		Yes		Yes	
Observations	1,501		734		767	
Model statistics	LL = −2,682.1		LL = −1,291.4		LL = −1,280.1	
	F-stat = 10.9 ($p = 0.00$)		F-stat = 5.3 ($p = 0.00$)		F-stat = 5.6 ($p = 0.00$)	

Note. Standard errors are based on state level clustering.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

is not required for generator adoption. Overall, this eliminates the possibility that literacy, corruption/enforcement, and electric power factors are all proxies for an unobservable development variable that might commonly affect all types of technology (computer and generator) adoption, as well as productivity. Thus we have greater faith in the transparency mechanism affecting computer adoption.

Furthermore, note that some transparency-related variables included in the regression are at the firm level, obtained through the survey (e.g., perceived informality by peers, external audit, etc.), whereas others are at the state level (e.g., TI corruption index, regulatory consistency). Neither the state level nor firm-specific transparency measures are significant in the generator adoption equation though they were significant in computer adoption. As transparency metrics at the state level are potentially correlated with other omitted factors such as lower levels of development and less established infrastructure in the state, it is gratifying that not only the state but also firm-specific (local) transparency factors are not significant.[12]

4.4. Impact of Computer Adoption on Productivity

We next report the results of the productivity regressions, beginning with ordinary least squares (OLS) regressions. Tables 4(a) and 4(b) report the results of the regression for labor (revenue/employee) and floor area productivity (revenue/square foot). For each productivity variable, we report the results for the full sample and for large and small firms.[13] We begin with the results for the full sample. Productivity is higher among firms adopting computers as reflected by the positive coefficient on computer adoption. Given that the dependent variable enters the regression equation in logs, the productivity multiplier on revenue per employee for firms adopting computers is 25% ($\exp(0.224) = 1.25$); on revenue per square foot it is 29% ($\exp(0.259)$).

[12] We thank an anonymous reviewer for highlighting the importance of finding insignificant relationships not only at the state but also at the firm (local) level. As an additional robustness check, we ran the probit regression of generator adoption with state-level fixed effects by including only the firm-level variables but excluding all of the state-level variables. The results are qualitatively identical to those in Table 3 for all of the firm-level variables in terms of their statistical significance levels and directionality. These results are available in the online appendix. See Table 8(b) for the simultaneous equations maximum likelihood estimation (MLE) results on computer adoption with state fixed effects.

[13] Our analysis indicated statistically significant heterogeneity in gains from computer adoption by firm size, but not across other firm characteristics. Hence we focus only on large and small firm differences in the rest of the paper.

Table 4(b) Floor Area Productivity—OLS Regression

	All stores		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE
Productivity-enhancing technology adoptions						
Computer	0.259**	0.106	0.266***	0.101	0.247**	0.102
Generator	0.213**	0.086	0.209**	0.086	0.220***	0.081
Store characteristics						
Store size	−0.057	0.047	−0.059**	0.027	−0.056**	0.026
Store age	0.082**	0.052	0.108**	0.055	0.078**	0.039
Store management and ownership characteristics						
Manager experience	0.093*	0.050	0.123**	0.057	0.087**	0.039
Ownership concentration	−0.003	0.002	−0.005*	0.003	−0.003	0.002
Government owned	0.492	0.316	0.532	0.348	0.382	0.285
Finance, in-store security, and competitive factors						
Bank account	0.298***	0.078	0.348***	0.090	0.288***	0.077
Overdraft facility	0.223***	0.086	0.242***	0.092	0.221***	0.082
In-store security	0.228***	0.082	0.245***	0.090	0.213***	0.076
Price competition level	0.104	0.070	0.094	0.069	0.108	0.078
Inventory level for main product	0.002	0.002	0.003*	0.001	0.003	0.003
State level educational factor						
Literacy rate	0.026**	0.012	0.027**	0.013	0.025**	0.012
Fixed effects						
Store type and city	Yes		Yes		Yes	
Observations	1,503		734		769	
Model statistics	LL = −2,607.8		LL = −1,680.9		LL = −1,342.4	
	F-stat = 10.4 (p = 0.00)		F-stat = 6.06 (p = 0.00)		F-stat = 6.36 (p = 0.00)	

Note. Standard errors are based on state level clustering.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

The other variables in the regression have the expected signs for both metrics of productivity. Firms that adopt generators have higher productivity. Larger stores are more productive. Interestingly, manager experience is positively related to store productivity. Thus, even though experience is negatively related to computer adoption and thus may be associated with lower productivity in the absence of computer adoption, the direct relationship between managerial experience and productivity is positive. Private sector ownership is negatively related to productivity. This may reflect the fact that much of single proprietor ownership is driven by subsistence stores. In terms of store characteristics, access to banking and financing is positively correlated with greater productivity. Use of in-store security is also positively correlated with productivity, as theft is widely considered a serious drain in retailing.

To eliminate spurious relationships between transparency variables and productivity, we have included a number of controls that may be correlated with transparency variables. Still, there may be other (possibly omitted) variables that could be correlated with the transparency variables, and which have their own direct effect on computer adoption, but not on generator adoption. Literacy rate is an example of such a variable. Demands from an educated workforce may reduce corruption and increase enforcement. This, in turn, can increase computer adoption, but not generator adoption. We therefore estimate the model including literacy rate in the productivity and computer/generator adoption regression. Literacy rates are significant in the computer adoption and productivity equation, but not in the generator adoption regression as noted earlier.[14] The fact that the productivity-enhancing effects of computers remains significant even after controlling for the effects of moderating variables such as literacy (or HDI) that are correlated not only with computer adoption and productivity but also with transparency variables, lends confidence to the conclusion.

As earlier, we report the results of the productivity regressions for large and small firms based on a median split. The direction of the estimates is qualitatively identical for large and small firms just as in the full sample estimates for labor and floor area productivity metrics. Comparing the results by firm size, larger firms gain more in terms of labor productivity from computer adoption (29%, i.e., exp(0.253)) relative to smaller firms (24%). The corresponding numbers for floor area productivity are 28% and 24%,

[14] The results are qualitatively identical when HDI is included in the regression in place of literacy rate.

respectively. Thus our results are consistent with the notion that productivity gains are larger for large stores that require greater coordination. The other variables in the regressions have the same signs as shown earlier.

4.5. Propensity Score Matching

One concern with the OLS estimates on the effect of computers on productivity is that stores that adopt computers are systematically different from stores that do not. Thus, differences in productivity across the two groups may not be due to computer adoption. Matching methods, pioneered by Rosenbaum and Rubin (1983) and refined by Heckman and colleagues (e.g., Heckman et al. 1997, 1998a, b), have been developed such that the outcomes of the treated (computer adopters) denoted by Y_1 are contrasted only against the outcomes of comparable untreated (nonadopters) denoted by Y_0. In this way productivity differences can be attributed to the treatment (computer adoption). The basic idea of the matching method is discussed below.

Let I_0 and I_1 denote the set of indices for nontreated and treated, respectively. To estimate a treatment effect for each treated firm $i \in I_1$, outcome Y_{1i} is compared against the average of outcomes Y_{0j} for all matched firms $j \in I_0$ among the untreated firms. Matches are based on observed characteristics that affect treatment (in our case, the variables that impact computer adoption as reported in Table 2). When the observed characteristics of the untreated firm are closer to that of the treated firm, based on an appropriate distance metric, the untreated firm is weighted higher when constructing the match. Thus the estimated gain for each firm i in the treated sample I_1 is $Y_{1i} - \sum_{j \in I_0 \cap S_p} W(i,j) Y_{0j}$, where S_p is the set of firms in the region of common support across the treated and nontreated, i.e., $S_p = \mathrm{Supp}(X \mid D=1) \cap \mathrm{Supp}(X \mid D=0)$ and $W(i,j)$ is an algorithm-specific weight based on the distance between the propensity scores for i and j. Let n_1 be the number of treated cases. The focal parameter of interest, called the Average Treatment effect on the Treated (ATT), reflecting the average effect of computer adoption on productivity is defined as

$$\frac{1}{n_1} \sum_{j \in I_0 \cap S_p} \left(Y_{1i} - \sum_{j \in I_0 \cap S_p} W(i,j) Y_{0j} \right).$$

Specifically, we use kernel matching advocated in Heckman et al. (1997) to construct the weighting function based on the difference in propensity scores between firm i and j. The weighting function is given by

$$W(i,j) = \frac{G((P_j - P_i)/h)}{\sum_{k \in I_0} G((P_k - P_i)/h)},$$

where $G(\cdot)$ is the kernel function, P_j is the propensity score of firm j, and h is a tuning parameter that specifies a bandwidth for the kernel function. Specifically, we report results based on the Epanechnikov kernel $G(u) = 0.75(1-u^2)I(|u| \le 1)$ (see Leuven and Sianesi 2003). The kernel has a parabolic shape with support in the region $[-1, 1]$.[15] We estimate the model using the PSMATCH2 module in Stata.

To perform PSM, a critical requirement is the ability to match treated observations with nontreated observations through propensity scores. For this, a rich set of variables that can reasonably discriminate the treated and nontreated observations is necessary. Specifically, we estimate the propensity scores through probit analysis of computer adoption using the same set of observed covariates that we report in the probit regressions shown in Table 2. Given missing data on different covariates, we have a sample size of 1,501 with 324 treated and 1,177 nontreated cases. The probit model provides a good fit with the data with an adjusted R^2 of 0.52. Given the requirements of common support for the estimated propensity scores between treated and nontreated firms, the propensity matching involves a sample of 1,199 with 269 treated and 930 nontreated units. The region of common support is $[0.004, 0.999]$ and the mean propensity score for computer adopters is 0.58. Table 5 shows the distribution of the estimated propensity scores and the mean values of selected (and representative of key dimensions) variables used in PSM for the treated and nontreated units by distinct block grouping.[16] Block grouping ensures that, within each block (PS interval), the mean values of estimated propensity scores are very comparable between treated and nontreated units. The number of blocks, here 13, is generated by enforcing that condition using the algorithm by Becker and Ichino (2002) based on repeated splitting of each block starting with 5 equal interval initial blocks until the comparable condition is achieved. As expected, there are more untreated firms in the low propensity score blocks, whereas there are more treated firms in the high propensity score blocks. This suggests that the variables included for propensity matching facilitate discrimination of the firms based on computer adoption.

Table 6 reports the ATT estimates based on labor and size productivity measures. We first report the results for the full sample. Our results show that

[15] We also estimated the effect using other matching methods such as nearest neighbor matching based on both propensity scores (Leuven and Sianesi 2003) as well as actor norm based distances between treated and control units (Abadie et al. 2004). Our results are similar in magnitude across the different methods.

[16] Results for all of the variables used in our analyses are given in the online appendix.

Table 5 Distribution of Propensity Scores and Means for Selected Propensity Scoring Variables Over the Common Support

		Number of sample observations		Means for selected variables in the treated and control groups[b]					
Block or PS interval #	Lower bound of propensity score within block[a]	Treated	Control	% revenue spent on bribe	Regulatory consistency (1–6)	Power supply problem (0–4)	Gross annual revenue (million Rs.)	Store size (sq. ft.)	Managerial experience (yrs.)
1	0.000	4	328	0.442; 0.461	3.237; 3.097	1.603; 1.568	0.625; 0.597	128.50; 142.00	11.00; 11.37
2	0.025	6	132	0.985; 1.004	3.133; 3.076	1.516; 1.619	0.925; 0.792	199.67; 182.08	11.83; 11.52
3	0.050	8	134	0.939; 0.980	3.129; 3.102	1.572; 1.641	1.229; 1.090	208.75; 217.23	11.75; 13.05
4	0.100	33	135	1.012; 1.019	3.124; 3.116	1.545; 1.610	1.741; 1.592	269.73; 247.90	12.61; 12.81
5	0.200	5	59	0.874; 0.815	2.900; 3.091	1.802; 1.601	1.840; 1.589	280.40; 318.10	13.80; 13.22
6	0.250	11	33	1.708; 1.663	3.108; 3.186	1.524; 1.615	2.564; 2.235	321.82; 332.88	13.27; 13.73
7	0.300	18	33	1.585; 1.630	2.991; 3.208	1.546; 1.644	4.706; 4.962	424.72; 411.39	14.22; 13.21
8	0.400	7	23	0.947; 0.996	3.236; 3.100	1.616; 1.732	6.357; 5.966	379.71; 331.74	10.14; 11.70
9	0.450	17	10	2.276; 2.295	2.823; 2.911	1.446; 1.536	5.956; 5.827	636.59; 611.00	14.94; 14.20
10	0.500	20	17	1.034; 0.932	2.968; 3.203	1.481; 1.659	5.150; 4.986	611.80; 583.53	14.45; 14.00
11	0.600	21	13	1.133; 1.112	3.115; 3.166	1.517; 1.364	5.426; 5.512	707.62; 705.38	12.24; 12.85
12	0.700	35	9	0.955; 0.867	3.144; 3.239	1.459; 1.604	5.533; 5.609	840.86; 912.11	11.51; 12.44
13	0.800	84	4	1.028; 0.946	3.141; 3.069	1.465; 1.358	7.489; 7.750	5,596.01; 5,250.00	13.50; 13.25

[a]The region of common support is [0.004, 0.999].
[b]For each selected variable, mean values are shown by treated and control groups (in that sequence) within each block.

Table 6 Propensity Score Analysis with Kernel Matching

Outcome	Sample	ATT[a]	S.E.
Log of labor productivity	All stores	0.409***	0.142
Log of floor area productivity	All stores	0.528***	0.187
Log of labor productivity	Large stores	0.418***	0.160
	Small stores	0.367**	0.158
Log of floor area productivity	Large stores	0.557***	0.128
	Small stores	0.527**	0.241

[a]All estimates are based on bias-corrected matching estimators using the kernel (Epanechnikov) matching approach (Leuven and Sianesi 2003).
$^{**}p < 0.05$; $^{***}p < 0.01$.

computer adoption enhances labor productivity by about 51% (ATT value of 0.409) and floor productivity by 70% (ATT value of 0.528). Clearly, the OLS results substantially underestimate the productivity increase from computer adoption. Similarly, the ATT for large and small firms reported also show that OLS estimates are substantially biased downwards. Overall, the qualitative insight that larger firms gain more in productivity from computer adoption continues to hold.

Question: Why are the OLS estimates on the productivity increase due to computer adoption biased downwards? At the margin, factors affecting computer adoption (including transparency) raise the threshold of productivity required for computer adoption. If we do not control for these factors, the threshold for productivity required for computer adoption is lower. Thus, the OLS estimates that do not account for the factors affecting adoption have a downward bias in productivity increases.

4.5.1. Selection on Unobservables: Rosenbaum Bounds. Within the PSM framework, sensitivity to potential selection on unobservables (hidden selection not captured in the observable variables used in PSM) is assessed using Rosenbaum bounds (Rosenbaum and Rubin 1983, DiPrete and Gangl 2004). Basically, we assess how much of the variance in unobservables needs to drive selection to negate the treatment effect. The higher the variance needed, the more confident we are that the qualitative results about the role of computer adoption are robust.

Table 7 shows the level of unobserved variance necessary to make the productivity enhancing effects of computers insignificant due to unobserved selection. The treatment effect becomes insignificant when $\Gamma > 1.9$ for the labor productivity measure and at about $\Gamma > 2.2$ for the floor productivity measure. In this context, the mean propensity score for computer adopters in our study is 0.58. Our findings of the treatment effect on the labor productivity measure remain robust to unobserved selection effects until about $\Gamma = 1.9$. This indicates that unobserved selection bias will undermine our finding of the positive productivity impact of computer adoption if the mean propensity of computer adopters increases from 0.58 to $0.58 \times 1.9 = 1.1$. Similarly, our findings of the treatment effect on the floor productivity measure remain robust to unobserved selection effects until about $\Gamma = 2.2$. This indicates that unobserved selection bias will undermine our finding of the positive productivity impact of computer adoption if the mean propensity of computer adopters increases from 0.58 to $0.58 \times 2.2 = 1.3$. Because it is highly unlikely that the probability of computer adopters will jump from 58%

Table 7 Robustness to Unobserved Selection Effects: Rosenbaum Bound Analyses

	p-Value[b]		H-L point estimate		Conf. interval[b]	
Gamma (Γ)	(U-Bound)	(L-Bound)	(U-Bound)	(L-Bound)	(U-Bound)	(L-Bound)
			For log of labor productivity[a]			
1.0	0.0000	0.0000	0.4502	0.4502	0.3163	0.5713
1.1	0.0000	0.0000	0.4068	0.4886	0.2751	0.6141
1.2	0.0000	0.0000	0.3688	0.5240	0.2371	0.6519
1.3	0.0001	0.0000	0.3329	0.5570	0.2036	0.6844
1.4	0.0001	0.0000	0.3017	0.5902	0.1720	0.7150
1.5	0.0001	0.0000	0.2718	0.6195	0.1426	0.7429
1.6	0.0001	0.0000	0.2458	0.6468	0.1133	0.7688
1.7	0.0006	0.0000	0.2211	0.6700	0.0839	0.7948
1.8	0.0099	0.0000	0.1975	0.6922	0.0582	0.8182
1.9	0.0263	0.0000	0.1742	0.7133	0.0327	0.8391
2.0	0.0581	0.0000	0.1527	0.7333	0.0100	0.8601
2.1	0.1103	0.0000	0.1307	0.7526	−0.0102	0.8797
2.2	0.1847	0.0000	0.1097	0.7717	−0.0303	0.8980
2.3	0.2787	0.0000	0.0902	0.7886	−0.0503	0.9163
2.4	0.3856	0.0000	0.0709	0.8044	−0.0681	0.9329
2.5	0.4968	0.0000	0.0541	0.8203	−0.0862	0.9479
			For log of floor area productivity[a]			
1.0	0.0000	0.0000	0.5919	0.5919	0.4645	0.7203
1.1	0.0000	0.0000	0.5410	0.6414	0.4120	0.7721
1.2	0.0000	0.0000	0.4950	0.6895	0.3626	0.8212
1.3	0.0000	0.0000	0.4523	0.7318	0.3166	0.8621
1.4	0.0000	0.0000	0.4116	0.7725	0.2748	0.9009
1.5	0.0000	0.0000	0.3732	0.8105	0.2324	0.9358
1.6	0.0001	0.0000	0.3382	0.8447	0.1959	0.9702
1.7	0.0004	0.0000	0.3041	0.8745	0.1617	1.0022
1.8	0.0013	0.0000	0.2729	0.9031	0.1289	1.0350
1.9	0.0037	0.0000	0.2399	0.9300	0.0991	1.0642
2.0	0.0091	0.0000	0.2112	0.9539	0.0709	1.0933
2.1	0.0195	0.0000	0.1857	0.9810	0.0419	1.1190
2.2	0.0375	0.0000	0.1601	1.0037	0.0139	1.1430
2.3	0.0654	0.0000	0.1355	1.0270	−0.0125	1.1662
2.4	0.1051	0.0000	0.1134	1.0497	−0.0397	1.1878
2.5	0.1572	0.0000	0.0919	1.0706	−0.0624	1.2116

Note. HL, Hodges-Lehmann.

[a]Results are based on differences between computer adopters and matched nonadopters, using kernel (Epanechnikov) matching on propensity scores through Leuven and Sianesi (2003).

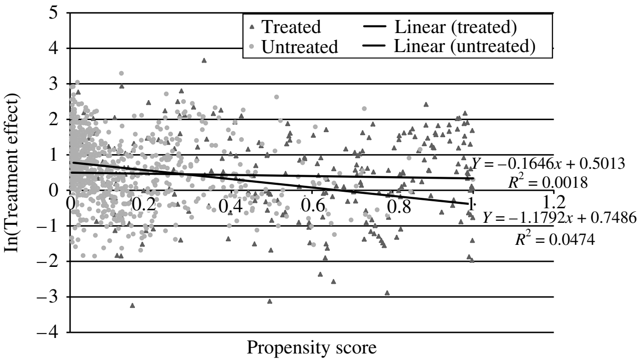
[b]*p*-values and confidence intervals are one-sided and at the 90% level.

(based on observables alone) to 110% or 130% (including unobservable effects), our PSM-based findings on the positive productivity impact of computer adoption are reasonably robust to hidden selection bias. This gives us faith in the finding that computer adoption enhances productivity.

4.5.2. Heterogeneous Treatment Effects. We next explore heterogeneous treatment effects across firms. We use the *Matched Smoothing Method of Estimating Heterogeneous Treatment Effects* (MS-HTE) described in Brand and Xie (2010) and Xie et al. (2011). The PSM results reported the ATT. As noted, for labor productivity, the ATT is 0.409. Thus computer adoption enhances labor productivity for the treated by 51%. The AT for the Untreated (ATU) is 0.605. Thus computer adoption enhances labor productivity for the untreated by 83%. The average treatment effect (across Treated and Untreated) is 0.561. Thus computer adoption enhances labor productivity on average by 75%. However there is heterogeneity in these effects.

Scatter plots of the estimated treatment effects against the propensity scores for the treated and untreated are shown in Figure 10. A regression curve fitted on these scatter plots shows that, overall, there is a decline in treatment effects for treated and untreated groups with increased propensity scores. However the slope is relatively flat and insignificant for the treated, while significantly negative for the untreated. More important, we find that when propensity scores are low, the productivity increases for untreated firms are greater than the productivity increases for treated firms. Yet when propensity scores are high, the productivity increases for untreated firms are lower than the productivity increases for treated firms. This is consistent with our earlier argument that, at the margin, factors affecting computer adoption

Figure 10 Heterogeneity in Treatment Effects



(including transparency) raise the productivity threshold required for computer adoption. Thus productivity increases for the untreated are higher when the propensity scores are lower. In spite of potentially high gains in productivity, other factors reduce the propensity to adopt computers. However when such impediments to adoption are lower (as when the propensity score is higher), productivity increases for treated firms are higher. On average, as there are far more untreated firms, and they tend to be at the low end of the propensity score spectrum, we find that ATU > ATT.

4.6. Modeling Self-Selection

With Rosenbaum bounds, we considered selection on unobservables as if the selection is random. Yet if selection is nonrandom, as is the case with computer adoption, one has to model self-selection. To address this concern, we next estimate the effect of computer adoption on firm productivity using Heckman's model of self-selection.[17]

Let y_i denote the outcome, i.e., productivity of firm i, and x_i be factors that impact productivity. Let w_i be the self-selected choice of firm i to adopt computers and z_i denote factors that affect the decision of firm i to adopt computers. Let w_i^* be a latent variable indicating the incremental value obtained by firm i by adopting computers, and $w_i = 1$ if $w_i^* > 0$ and $w_i = 0$ otherwise. The Heckman SSM is described by the following outcome and selection equations, respectively:

$$y_i = x_i\beta + \alpha w_i + \varepsilon_i,$$
$$w_i^* = z_i\gamma + \nu_i,$$

where

$$\begin{pmatrix} \varepsilon_i \\ \nu_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix}\right].$$

The bivariate normal model above can be estimated by Heckman's two-step estimation approach or maximum likelihood (Maddala 1983). We estimate the model using maximum likelihood. Note that revenue is an endogenous variable in the computer adoption equation, and computer adoption is an endogenous variable in the revenue equation. As there are many exogenous variables that are only present in the revenue or the computer adoption equation, and not in both, the system is identified.

We report the estimation results in Table 8. Table 8(a) reports the results of the labor productivity equation. Table 8(b) reports the results of the computer adoption selection equation. Table 8(c) reports the results of the floor area productivity equation. We suppress the results of the computer adoption selection equation corresponding to floor area productivity as these results are virtually identical to Table 8(b). We run each of the SSMs in Tables 8(a)–8(c) with and without the state-level fixed effects. The models with state-level fixed effects include only the firm- level variables but exclude all of the state-level variables. This allows us to rule out any kind of potential state-level unobservables. As a comparison of the results across the columns in Tables 8(a)–8(c) shows, they remain virtually identical with or without controlling for the state-level fixed effects.[18] Also, the signs of the variables in productivity equations in Tables 8(a) and 8(c) remain essentially the same as in the OLS results reported in Tables 4(a) and 4(b). However, there is considerable bias in the OLS estimates of computer adoption on productivity. Of particular importance, the correlation between the outcome and selection term is negative and significant with values between −0.42 to −0.53 across the different models, suggesting the importance of unobserved selection.

Using the full sample SSM for labor productivity, we elaborate on the estimated average marginal effects of a few variables of interest that indirectly impact productivity through their effect on computer adoption.[19] For example, an external auditor for a retailer

[17] The PSM literature argues that the Rosenbaum bounds approach does not require normal distribution assumptions and is nonparametric, unlike the bivariate normal Heckman selection model. Another advantage is that the Heckman selection model requires exclusion restrictions involving instruments, i.e., variables that impact selection, but not outcomes. These may not always be available. Furthermore, these estimates are not the average treatment effect across all firms, but a local average treatment effect (LATE) over the firms whose decision to adopt computers is affected by the instruments. When selection on unobservables is not random, and exclusion restrictions are available, the selection model is preferred.

[18] To conserve space, we have reported here the results of SSMs with state fixed effects only for the full sample. Corresponding results by store size are qualitatively identical and available in the online appendix.

[19] We use the margins command in Stata to compute the average marginal effects reported.

Table 8(a) Labor Productivity Outcome Equation—MLE for Self-Selection Model

	All stores-I		All stores-II		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Productivity enhancing technology adoptions								
Computer	0.464***	0.066	0.471***	0.074	0.517***	0.142	0.454***	0.206
Generator	0.524**	0.237	0.537**	0.224	0.473**	0.216	0.544**	0.267
Store characteristics								
Store size	0.223***	0.075	0.224***	0.076	0.230**	0.095	0.213***	0.074
Store age	−0.065	0.141	−0.082	0.124	−0.094	−0.066	−0.075	0.061
Store management and ownership characteristics								
Manager experience	0.126***	0.046	0.135***	0.046	0.138***	0.060	0.108***	0.041
Ownership concentration	−0.006***	0.002	−0.006***	0.002	−0.006**	−0.003	−0.007**	0.003
Government owned	0.475	0.319	0.465	0.310	0.686	0.501	0.385	0.335
Finance, in-store security, and competitive factors								
Bank account	0.306***	0.079	0.317***	0.066	0.368***	0.097	0.283**	0.118
Overdraft facility	0.407**	0.196	0.389**	0.183	0.415**	0.187	0.370**	0.176
In-store security	0.328**	0.151	0.334**	0.151	0.343***	0.097	0.326***	0.122
Price competition level	0.077	0.066	0.077	0.059	0.122	0.088	0.067	0.114
Inventory level for main product	0.005***	0.002	0.005***	0.002	0.008***	0.003	0.005***	0.002
State level educational factor								
Literacy rate	—	—	0.019*	0.010	0.020*	0.010	0.020**	0.009
Fixed effects	Store type and state		Store type and city		Store type and city		Store type and city	
Observations	1,501		1,501		734		767	
Model statistics[a]	LL = −2,566.28 $\chi^2 = 524.7$ ($p = 0.00$) $\rho = -0.44$		LL = −2,494.92 $\chi^2 = 660.6$ ($p = 0.00$) $\rho = -0.46$		LL = −1,272.86 $\chi^2 = 550.2$ ($p = 0.00$) $\rho = -0.51$		LL = −1,350.32 $\chi^2 = 400.6$ ($p = 0.00$) $\rho = -0.42$	

Note. Standard errors are based on clustering at city level for models with state fixed effects, and at state level for models with no state fixed effects.
[a] ρ denotes the correlation in error terms between the outcome (productivity) and the selection (computer adoption) equations.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

is related to 6.98% higher probability of computer adoption, which in turn translates to a 24.11% higher labor productivity based on simultaneous estimation of productivity and selection model. Similarly a standard deviation increase in the regulatory inspections and regulatory consistency variables is related to a 9.20% and 5.02% higher probability of computer adoption, which in turn translate to 11.75% and 4.57% higher labor productivity. Finally, we consider managerial experience, which affects both computer adoption and productivity. The main effect of managerial experience on store productivity is positive. The moderating effect of managerial experience on productivity through computer adoption is negative. The net effect of a standard deviation increase in managerial experience on productivity is higher by 17.84%, after accounting for the negative effect on computer adoption.

To facilitate comparison across the different estimation approaches, the estimated effects of computer adoption on productivity using OLS, PSM, and SSM are reported in Table 9. The estimates using SSM are close to those from the PSM but substantially greater than those from the OLS. To be specific, in terms of labor productivity, while OLS estimates a 25% improvement, PSM estimates a 50% improvement, and SSM estimates a 60% improvement on the treated firms. The heterogeneous treatment effects model estimates the average treatment effect on the untreated at 83% and the average treatment effect across treated and untreated at 75%. Across all of the results, it is clear that not accounting for transparency leads to a significant underestimation of productivity improvement from computer adoption. The results are directionally and substantively consistent for floor area productivity and also across small and large firms.

5. Conclusion

The tendency of firms in emerging markets to avoid productivity-enhancing technologies and remain small due to transparency concerns has been dubbed the Peter Pan Syndrome. Though IT enhances productivity, the culture of informality in emerging markets causes businesses to fear IT because it removes the veil of secrecy that is conducive to tax evasion. This paper investigated whether emerging market firms make the trade-off between productivity and transparency in adopting IT.

Specifically, we studied the productivity-transparency trade-off in the Indian retail sector using data from

Table 8(b) **Business Computer Adoption Selection Equation—MLE for Self-Selection Model**

	All stores-I		All stores-II		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Corruption factors discouraging transparency								
State TI Corruption Index	—	—	−0.110***	0.036	−0.116***	0.037	−0.093***	0.028
Percent revenue spent on bribe	−0.028*	0.014	−0.024*	0.012	−0.026**	0.013	−0.019*	0.011
Perceived informality by peers	−0.004***	0.001	−0.004***	0.001	−0.004***	0.002	−0.002***	0.001
Enforcement factors encouraging transparency								
Regulatory inspections	0.040***	0.016	0.041***	0.013	0.048***	0.014	0.025***	0.008
External auditor	0.274***	0.095	0.278***	0.099	0.236***	0.082	0.304***	0.104
Regulatory consistency	—	—	0.025**	0.012	0.024**	0.012	0.026**	0.013
Labor cost and education level								
State minimum wage	—	—	0.021**	0.011	0.022**	0.011	0.015**	0.007
State literacy rate	—	—	0.065*	0.036	0.067*	0.036	0.062**	0.029
Electric power supply related factors								
Power outage	−0.063	0.102	−0.067	0.117	−0.077	0.108	−0.063	0.098
State power supply problem	—	—	−0.081	0.056	−0.084	0.059	−0.074	0.061
Productivity enhancing technology adoptions								
Generator	0.406***	0.125	0.398***	0.117	0.491***	0.139	0.519***	0.148
Store level performance measure								
Gross annual revenue	0.093***	0.008	0.095***	0.007	0.087***	0.006	0.133***	0.014
Store characteristics								
Store size	0.291***	0.063	0.299***	0.064	0.274***	0.055	0.353**	0.168
Employee size	0.239***	0.038	0.235***	0.038	0.268***	0.041	0.196***	0.019
Store age	−0.086**	0.042	−0.091**	0.045	−0.108**	0.051	−0.088**	0.042
Store management and ownership characteristics								
Manager experience	−0.146*	0.070	−0.140*	0.075	−0.177**	0.089	−0.127*	0.068
Ownership concentration	−0.009	0.006	−0.011*	0.006	−0.010*	0.006	−0.008*	0.005
Government owned	−0.322	0.223	−0.336	0.237	−0.532	0.409	−0.197	0.768
Fixed effects	Store type and state		Store type and city		Store type and city		Store type and city	
Observations	1,501		1,501		734		767	
Model statistics	LL = −2,566.09 $\chi^2 = 524.8$ ($p = 0.00$)		LL = −2,494.94 $\chi^2 = 660.6$ ($p = 0.00$)		LL = −1,272.88 $\chi^2 = 550.2$ ($p = 0.00$)		LL = −1,350.33 $\chi^2 = 400.6$ ($p = 0.00$)	

Note. Standard errors are based on clustering at city level for models with state fixed effects and at state level for models with no state fixed effects.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

a large scale national survey of 1,948 Indian retailers augmented with other relevant data on corruption, enforcement, and other state-level control variables. We find that IT adoption is significantly affected by transparency concerns. Whereas corruption reduces IT adoption, enforcement and auditing increases IT adoption by providing a level playing field for all firms and reducing the negative impact of corruption. IT adoption increases store productivity on average by about 50% to 70%. The effects of transparency on IT adoption and the impact of adoption on productivity are both greater for larger than smaller firms. At the margin, higher corruption and lower enforcement raises the threshold of productivity required for IT adoption.

Our results are relevant to transparency-enhancing IT businesses, government, and policy makers. As growth in the developed world stagnates, firms increasingly rely on emerging markets for their growth. To the extent that the market potential for IT among businesses is linked to enhanced productivity, our results show that corruption and enforcement levels in a market impact not only unit sales but also the willingness to pay (and therefore the price) in emerging markets. For government and policy makers, our results suggest that forceful enforcement and corruption reduction can not only have a direct positive impact on tax collection but also an indirect positive impact on the tax revenue base. The latter occurs through greater productivity induced by the use of modern efficiency-enhancing technologies and by bringing more businesses into the transparent formal sector. Our work shows that modeling the institutional characteristics of emerging markets can enhance the relevance of academic research for managers and policy makers.

We conclude with a discussion of the limitations of the paper, which provide possibilities for future

Table 8(c) Floor Area Productivity Outcome Equation—MLE for Self-Selection Model

	All stores-I		All stores-II		Larger stores		Smaller stores	
Variables	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Productivity enhancing technology adoptions								
Computer	0.538***	0.105	0.543***	0.103	0.566***	0.056	0.523***	0.082
Generator	0.410**	0.178	0.418**	0.188	0.394**	0.163	0.440**	0.195
Store characteristics								
Store size	−0.065***	0.015	−0.061***	0.012	−0.066***	0.022	−0.059***	0.010
Store age	0.142**	0.063	0.147**	0.063	0.204***	0.077	0.143**	0.070
Store management and ownership characteristics								
Manager experience	0.175**	0.076	0.176**	0.076	0.214***	0.070	0.165***	0.052
Ownership concentration	−0.005**	0.002	−0.005**	0.002	−0.007***	0.003	−0.005*	0.003
Government owned	0.361	0.356	0.337	0.343	0.307	0.394	0.363	0.272
Finance, in-store security, and competitive factors								
Bank account	0.342***	0.083	0.335***	0.082	0.448***	0.116	0.314**	0.138
Overdraft facility	0.143*	0.075	0.137*	0.073	0.161*	0.091	0.059*	0.034
In-store security	0.206***	0.078	0.208***	0.076	0.251***	0.095	0.204**	0.092
Price competition level	0.074	0.067	0.076	0.065	0.076	0.087	0.000	0.089
Inventory level for main product	0.003**	0.002	0.003**	0.002	0.004**	0.002	0.003**	0.001
State level educational factor								
Literacy rate	—	—	0.022**	0.011	0.025*	0.013	0.023**	0.010
Fixed effects	Store type and state		Store type and city		Store type and city		Store type and city	
Observations	1,501		1,501		734		767	
Model statistics[a]	LL = −2,561.36 $\chi^2 = 710.83$ ($p = 0.00$) $\rho = -0.46$		LL = −2,595.52 $\chi^2 = 739.04$ ($p = 0.00$) $\rho = -0.49$		LL = −1,362.07 $\chi^2 = 670.02$ ($p = 0.00$) $\rho = -0.53$		LL = −1,452.27 $\chi^2 = 450.69$ ($p = 0.00$) $\rho = -0.46$	

Notes. Standard errors are based on clustering at city level for models with state fixed effects and at state level for models with no state fixed effects. Results for the computer adoption equation with the floor area productivity outcome equation are not shown to conserve space; the estimates are virtually identical to those presented in Table 8(b).

[a] ρ denotes the correlation in error terms between the productivity and selection regression equations, respectively.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

research. First, the India findings need to be replicated in other emerging markets. Second, the results are based on a cross-sectional data set. We used a variety of statistical tools to make appropriate inference using cross-sectional data, e.g., PSM, heterogeneous treatment effects, and instrumental variable methods that account for self-selection to measure the productivity effects of IT. We assessed whether the effects of transparency on IT adoption are robust through a falsification test and allowing for state-level fixed effects to account for other potentially omitted state-level factors that may drive IT adoption. Future research should replicate our key findings with panel data, ideally with some form of experimental or quasiexperimental transparency variation due to changes in policy or regulations.

Even though we did not find significant heterogeneity in gains from computer adoption beyond firm size, it would be useful to explore firm characteristics that can drive differences in gains from computer adoption. Finally, we modeled computer adoption as a discrete variable in assessing productivity. Future research should focus on assessing the effect of IT spending rather than IT adoption as a discrete variable. It would also be of interest to understand the impact of IT investments on retail prices as IT lowers marginal costs but also increases fixed costs. We hope that our study will stimulate further academic research on these important research questions.

Table 9 Effect of Computer Adoption on Productivity: Summary[a]

	OLS		Propensity score matching		Self-selection model	
Firm size	Labor	Floor	Labor	Floor	Labor	Floor
All	1.251	1.296	1.505	1.696	1.602	1.721
Large	1.288	1.305	1.519	1.745	1.677	1.761
Small	1.241	1.280	1.443	1.694	1.575	1.687

[a]The productivity effects are obtained by taking the exponential of the parameter estimates in the respective tables because productivity enters in logs as dependent variables in the regression. For example, consider the effects on labor productivity from all store analysis results. For OLS, from Table 4(a), exp(0.224) = 1.251; for PSM from Table 6, exp(0.409) = 1.505; and for SSM from Table 8(a) column 2, exp(0.471) = 1.602.

In conclusion, note that, given the strong gains in productivity from IT adoption, we agree (albeit partly) with Sunder (2012) that "Indian retailers can and should break out of the self-defeating confines of the beliefs about the profitability of tax evasion" thus avoiding the "informality trap of lower productivity." Yet curing the Peter Pan Syndrome among Indian retailers would require the government to improve the business environment to be free from corruption, and enhance the level and consistency of enforcement. As India opens its markets to multinational, multibrand retailers, the need to increase productivity becomes even greater for domestic retailer survival. We hope our work encourages greater investment in productivity-enhancing technologies by Indian retailers, as they prepare themselves for new levels of competition (Reardon and Gulati 2008).

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2015.0921.

Acknowledgments

The authors thank Mushfiq Mobarak, Juanjuan Zhang, participants at the 2012 Marketing Science Conference in Boston, the 2012 Marketing Science Emerging Markets Conference at the Wharton Business School, the 2013 China India Insights Conference in New York, the 2014 NBER Summer Workshop on the Economics of Information Technology and Digitization, the 2015 AIM-AMA Sheth Doctoral Consortium in Dubai, and the marketing workshops at the University of Connecticut, University of Massachusetts, and University of Toronto for their helpful comments.

References

Abadie A, Drukker D, Herr JL, Imbens GW (2004) Implementing matching estimators for average treatment effects in Stata. *Stata J.* 4(3):290–311.

Amin M (2010) Computer usage and labor regulations in India's retail sector. *J. Development Stud.* 46(9):1572–1592.

Angelini P, Generale A (2008) On the evolution of firm size distributions. *Amer. Econom. Rev.* 98(1):426–438.

Azuma Y, Grossman H (2008) A theory of the informal sector. *Econom. Politics* 20(1):62–79.

Bartel A, Ichniowski C, Shaw K (2005) How does information technology really affect productivity? Plant-level comparisons of product innovation, process improvement and worker skills. *Quart. J. Econom.* 122(4):1721–1758.

Barua A, Kriebel CH, Mukhopadhyay T (1995) Information technology and business value: An analytic and empirical investigation. *Inform. Systems Res.* 6(1):3–23.

Basu S, Fernald JG, Oulton N, Srinivasan S (2003) The case of the missing productivity growth: Or, does information technology explain why productivity accelerated in the United States but not in the United Kingdom? Working Paper Series WP-03-08, Federal Reserve Bank of Chicago, Chicago.

Becker SO, Ichino A (2002) Estimation of average treatment effects based on propensity scores. *Stata J.* 2(4):358–377.

Bird RM, Zolt EM (2008) Technology and taxation in developing countries: From hand to mouse. *National Tax J.* LXI(4, Part 2): 791–821.

Black S, Lynch L (2001) How to compete: The impact of workplace practices and information technology on productivity. *Rev. Econom. Statist.* 83(3):434–445.

Brand JE, Xie Y (2010) Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *Amer. Sociol. Rev.* 75(2):273–302.

Brynjolfsson E, Hitt L (1996) Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Sci.* 42(4):541–558.

Chopra A (2011) India's push to tame tax evasion. *National* (June 18). Accessed July 20, 2012, http://www.thenational.ae/business/industry-insights/economics/indias-push-to-tame-tax-evasion.

Commander S, Harrison R, Menezes-Filho N (2011) ICT and productivity in developing countries: New firm-level evidence from Brazil and India. *Rev. Econom. Statist.* 93(2):528–541.

Cull R, Xu LC (2005) Institutions, ownership, and finance: The determinants of profit reinvestment among Chinese firms. *J. Financial Econom.* 77(1):117–146.

Dabla-Norris E, Gradstein M, Inchauste G (2008) What causes firms to hide output? The determinants of informality. *J. Development Econom.* 85(1/2):1–27.

Dasgupta S, Sarkis J, Talluri S (1999) Influence of information technology investment on firm productivity: A cross-sectional study. *Logist. Inform. Management* 12(1/2):120–129.

DeLone WH (1981) Firm size and characteristics of computer use. *MIS Quart.* 5(4):65–77.

Dewett T, Jones GR (2001) The role of information technology in the organization: A review, model, and assessment. *J. Management* 27(3):313–346.

Dhara T, Thomas C (2011) In India, tax evasion is a national sport. *Bloomberg Businessweek* (July 28). Accessed July 20, 2012, http://www.bloomberg.com/bw/magazine/in-india-tax-evasion-is-a-national-sport-07282011.html.

DiPrete T, Gangl M (2004) Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociol. Methodology* 34(1):271–310.

Fabiani S, Schivardi F, Trento S (2005) ICT adoption in Italian manufacturing: Firm-level evidence. *Indust. Corporate Change* 14(2):225–249.

Farrell D (2004) Boost growth by reducing the informal economy. *Asian Wall Street J.* (October 18). Accessed April 24, 2012, http://www.mckinsey.com/insights/mgi/in_the_news/boost_growth_by_reducing_the_informal_economy.

Foster L, Haltiwanger J, Krizan CJ (2002) The link between aggregate and micro productivity growth: Evidence from retail trade. NBER Working Paper 9120.

Friedman E, Johnson S, Kaufman D, Zoido-Lobaton P (2000) Dodging the grabbing hand: The determinants of unofficial activity in 69 countries. *J. Public Econom.* 76(3):459–492.

Gatti R, Honoratti M (2008) Informality among formal firms: Firm-level, cross-country evidence on tax compliance and access to credit. Policy Research Working Paper 4476, World Bank, Washington, DC.

Government of India (2011) *India Human Development Report 2011: Towards Social Inclusion*. Planning Commission (Oxford University Press, New Delhi, India).

Harris SE, Katz JL (1991) Firm size and the information technology investment intensity of life insurers. *MIS Quart.* 15(3):333–352.

Heckman J, Ichimura H, Todd P (1997) Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Rev. Econom. Stud.* 64(4):605–654.

Heckman JJ, Ichimura H, Todd P (1998a) Matching as an econometric evaluation estimator. *Rev. Econom. Stud.* 65(2):261–294.

Heckman J, Ichimura H, Smith J, Todd P (1998b) Characterizing selection bias using experimental data. *Econometrica* 66(5): 1017–1098.

Ichniowski C, Shaw K, Prennushi G (1997) The effects of human resource management practices on productivity. *Amer. Econom. Rev.* 87(3):291–313.

International Tax Compact (2010) Avoiding tax evasion and tax avoidance in developing countries. *German Federal Ministry for Economic Cooperation and Development* (BMZ, Eschborn, Germany).

Johnson S, Kaufman D, McMillan J, Woodruff C (2000) What do firms hide? Bribes and unofficial activity after communism. *J. Public Econom.* 76(3):495–520.

Jorgenson D (2001) Information technology and the US economy. *Amer. Econom. Rev.* 91(1):1–32.

Lefebvre LA, Lefebvre É, Elia E, Boeck H (2005) Exploring B-to-B e-commerce adoption trajectories in manufacturing SMEs. *Technovation* 25(12):1443–1456.

Leuven E, Sianesi B (2003) PSMATCH2—Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Software. Accessed June 14, 2012, http://ideas.repec.org/c/boc/bocode/s432001.html.

Love PE, Irani Z, Standing C, Lin C, Burn JM (2005) The enigma of evaluation: Benefits, costs and risks of IT in Australian small–medium-sized enterprises. *Inform. Management* 42(7):947–964.

Loveman GW (1994) An assessment of the productivity impact on information technologies. Allen TJ, Scott Morton MS, eds. *Information Technology and the Corporation of the* 1990*s*: *Research Studies* (MIT Press, Cambridge, MA), 84–110.

Maddala GS (1983) *Limited-Dependent and Qualitative Variables in Econometrics* (Cambridge University Press, Cambridge, UK).

Marcouiller D, Young LL (1995) The black hole of graft: The predatory state and the informal economy. *Amer. Econom. Rev.* 85(3):630–646.

Mesfin M (2012) Ethiopia: Tax authority unhappy with lack of cash register servicing. *Addis Fortune* (March 4). Accessed August 3, 2012, http://allafrica.com/stories/201203060697.html.

Miller AR, Tucker CE (2014) Electronic discovery and the adoption of information technology. *J. Law, Econom., Organ.* 30(2):217–243.

Mishra P, Subramanian A, Topalova P (2008) Tariffs, enforcement, and customs evasion: Evidence from India. *J. Public Econom.* 92(10–11):1907–1925.

Morgan A, Colebourne D, Thomas B (2006) The development of ICT advisors for SME businesses: An innovative approach. *Technovation* 26(8):980–987.

Morrison CJ, Berndt ER (1990) Assessing the productivity of information technology equipment in the U.S. manufacturing industries. NBER Working Paper 3582.

Nayar K (2011) Laundering black money. *Deccan Herald* (February 4). Accessed August 17, 2013, http://www.deccanherald.com/content/134580/laundering-black-money.html.

People's Daily (2000) China battles tax evasion with help of computers. (December 4). Accessed August 17, 2013, http://english.peopledaily.com.cn.

Rao VV (2010) Black, bold and bountiful. *Hindu Bus. Line* (August 13). Accessed July 15, 2013, http://www.thehindubusinessline.com/todays-paper/black-bold-and-bountiful/article1001195.ece.

Reardon T, Gulati A (2008) The supermarket revolution in developing countries. *Internat. Food Policy Res. Institute* (*IFPRI*), Policy Brief 2, June. Accessed July 11, 2012, http://www.ifpri.org/publication/supermarket-revolution-developing-countries.

Rosenbaum P, Rubin D (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55.

Russell B (2010) Revenue administration: Managing the shadow economy. Technical Notes and Manuals 10/14, International Monetary Fund, Washington, DC.

Schneider F, Enste D (2002) Shadow economies: Size, causes and consequences. *J. Econom. Literature* 38(1):77–114.

Sinha B (2003) Pan masala unit raided for Rs. 6-cr sales tax evasion. *Times India* (October 23). Accessed September 23, 2013, http://articles.timesofindia.indiatimes.com/.

Smith MG (2013) Doing it under the table: Unreported sales in India's manufacturing sector. Working paper, U.S. Department of Treasury, Washington, DC, http://ssrn.com/abstract=220822.

Stiroh K (2002) Information technology and the US productivity revival: What do the industry data say? *Amer. Econom. Rev.* 92(5):1559–1576.

Stiroh K (2010) Reassessing the role of IT in the production function: A meta-analysis. NBER Working Paper 12245.

Sunder S (2012) India's retarded retail sector. *Live Mint* (January 8). Accessed February 11, 2012, http://www.livemint.com/Opinion/v68no3uwkCrdJgdS9OPxrI/India8217s-retarded-retail-sector.html.

The Daily Star (2007) Electronic cash registers to be made mandatory (December 17). Accessed September 23, 2013, http://archive.thedailystar.net/newDesign/news-details.php?nid=15744.

Thong J (1999) An integrated model of information systems adoption in small businesses. *J. Management Inform. Systems* 15(4):187–214.

Transparency International (2012) Corruption Perceptions Index 2012. Accessed June 2, 2012, http://www.transparency.org/poli/cpi/.

Xie Y, Brand JE, Jann B (2011) Estimating heterogeneous treatment effects with observational data. *Research Report II-*729 (*February*) (Population Studies Center, Institute of Social Research, University of Michigan, Ann Arbor).