The Decision to Innovate: Aggregate Implications of Size-Based Distortions

Tyler C. Schipper *
Department of Economics, University of Oregon
November 18, 2014

Abstract

In this paper I investigate how size-based policy distortions affect firm level decisions in the context of a partial-equilibrium model. Aggregate productivity is improved through heterogenous firms making investments in innovation. I calibrate and simulate economies under different tax distortions, focusing on tax schemes that fluctuate based on firm size. I find that under some size-based distortions, aggregate output is similar to aggregate output in an undistorted economy. However, firms pursue less innovation, and aggregate total factor productivity is lower. Output per worker is considerably lower, echoing previous work that underscores the importance of TFP in explaining cross-country differences in income.

Keywords: Innovation; Firm Size; TFP

JEL Classification Numbers: O31; O47; L11

*Address: 1285 University of Oregon, Eugene, OR 97403, email: schipper@uoregon.edu. The author thanks Shankha Chakraborty and Nicholas Sly for comments and suggestions. Any remaining errors are, of course, my own.
1 Introduction

Differences in technology and total factor productivity (TFP) are familiar refrains for explaining cross-country differences in income. Both exogenous and endogenous growth models accept the importance of technology and innovation, yet there still remain questions about why technology improves productivity so differently across countries.\(^1\) Recent work investigating the extent of these differences has focused on the role of macroeconomic distortions, which cause inefficiencies in the allocation of resources.\(^2\) Thus far, no work has explicitly linked the process of technological change that is vital in determining aggregate TFP with the macroeconomic distortions that have been studied extensively in the literature. This chapter seeks to connect technological change with size-based distortions to study aggregate outcomes.

Early work on explaining TFP changes has been directed in two principal directions. Howitt (2000) and Klenow and Rodríguez-Clare (2005) document how TFP differences can arise from differences (in growth rates and levels, respectively) in innovation processes across countries. Alternatively, Restuccia and Rogerson (2008) began an important second avenue of research on TFP differences.\(^3\) Specifically, they look at misallocations within an economy that may be unfortunate results of policy. These two areas of investigation, and the host of papers that have followed, need to be connected.

The primary research question of this chapter is how size-based policy distortions

---

\(^1\)See Solow (1956) and Romer (1990) for seminal works.

\(^2\)For instance Hseih and Klenow (2009) and Alfaro et al. (2008).

\(^3\)Guner, Ventura, and Yi (2006) pursue a similar point by looking at how firm-size distributions are affected by laws that place limitations on the size of establishments, mostly in the form of taxes or subsidies based on firm size.
affect a firm’s innovation decisions and the aggregate effects of those decisions. To answer this question, it is important to have a model that includes innovation along with heterogeneous firms that are affected differently by the imposed distortions. Innovation in the model should be thought of as more general than the process of invention. Productivity changes occur both through innovation as well as adoption. As an instructive example, firms can improve their productivity by adopting just-in-time inventory management just like they can from inventing new machinery to build their product. Since both of these changes may enhance productivity, or decrease it if they are unsuccessful, I consider both to be innovation in the model.

A central component of this model, and this line of research generally, is that there exist distortions that vary based on firm size. To be clear, distortions throughout this chapter refer to output distortions, not input distortions. Taxes are levied on firms’ output rather than causing differences in input prices among firms. While the model developed subsequently will be used to test a variety of tax schemes, the most interesting scheme is motivated by Tybout (2000) in his survey of manufacturing firms in developing countries. He explains that small firms may be able to avoid costly regulations, corruption, and other costs of operation in lesser developed countries. On the other hand, large firms may benefit from these regulations and/or be able to influence them through the political process. It is the firms in the middle, that become just large enough to invite additional attention without reaping the benefits of being large, that are the least well-off. To mimic this story, I test a size-based distortion that takes a quadratic form. Both small and large firms face lower taxes than medium-sized firms.

I find that taxation schemes that are independent of firm size have predictable

---

4While empirical evidence is scarce, Goyette (2009) documents audit probabilities (effective tax rates) that vary with firm size.
effects such as decreasing average firm size, innovation, and output growth. Size-dependent tax policies have a significant effect on the size-distribution of firms, but total output seems to be affected very little by a tax schedule that reflects Tybout (2000). Evidence suggests that firms substitute investment in innovation for labor resources to avoid increased taxes. I find that under the current calibration, a Tybout-style tax lowers aggregate TFP even though total output remains relatively the same.

The research question naturally links the endogenous growth work with the more recent emphasis on national policy distortions focused on firm size. This link is important on several fronts. The innovation process itself offers little in terms of specific policy proposals and testable implications at the national level. Additionally, current work on firm-size distortions does not focus on the mechanism through which technology is accumulated. Bridging these areas can help us understand the effects of specific national distortions on the process of innovation and technology accumulation.

I use a simple framework in which firms are capable of innovating and raising their productivity in future periods. Firms are heterogeneous in their initial productivity as in Melitz (2003). Thus, the model generates firm-size distributions where firm size is correlated with, but not determined by, the productivity level of firms. Different distortions (implicit tax schemes) change the incentives for innovation and the resulting firm-size distributions.

There is a well-developed literature documenting the tax schedule and regulatory environments across the distribution of firm sizes. Unfortunately, these studies have usually focused on segments (i.e. large firms or small firms) of the distribution. The benefits of being a small firm are well documented. The seminal work of Rauch (1991) shows that there is a “missing middle” in the distribution of employment because
small firms did not want to become large enough to attract attention. By staying small, and informal, firms avoid costly regulation and corruption. Qualitatively, de Soto (1989) tells the same story in Peru where entrepreneurs sought to stay small to avoid additional costs of being formally recognized and regulated.

Newer work has underscored the theory and observations of Rauch (1991) and de Soto (1989) with quantitative results. Gallipoli and Goyette (2011) show that the audit probability in Uganda is based on firm size (number of workers) rather than any level of capital or revenue. Specifically, they find that the probability of being audited increases substantially for firms with more than 30 employees. Importantly, none of the works above discuss the potential benefits of being a large firm. In order to generate a “missing middle,” there must be incentives to, at some point, grow.

This incentive is documented by Levenson and Maloney (1997) who find that large firms often benefit from regulatory regimes. If regulation poses some form of fixed cost, increasing firm size decreases average costs. Further, the literature on lesser-developed countries is rife with examples of how political power and leverage are endowed to large firms. For example, Faccio (2006) documents the widespread nature of political connections and the strong correlation between firm size and political connections.

Certainly, some distortions may not fit well into an implicit tax rate as done in this work. For instance, Garcia-Santana and Pijoan-Mas (2014) evaluate the effects of the Small Scale Reservation Laws in India. These laws reserve certain goods for production by small firms. While these laws may not be best modeled as a simple distortionary tax as done here, they certainly represent a policy that is size-dependent.

5 The missing middle refers to employment as a percentage of total workers in medium-sized firms being less than employment in small and large firms, creating a missing middle. See Tybout (2000) for a more extensive discussion.
and is consistent with the story presented by Tybout (2000) in that firms prefer to stay small in order to produce certain goods.

Unfortunately, no work documents the actual implicit taxes faced by firms over the entire distribution of firms. The difficulty in doing so is two-fold. First, the data on taxes and regulations that firms face is difficult to collect over the entire distribution of firms, especially for small informal firms. The data that has been collected is mostly survey data that is suggestive of the types of patterns discussed above, but it is extremely difficult to work with quantitatively due to the subjective nature of the responses. Second, while authors like García-Santana and Pigoan-Mas (2014) and Gallipoli and Goyette (2011) look at the effects of specific policies that vary across firms, they do not capture the aggregated distortions that are found to exist in Hsieh and Klenow (2009). Firm-size distributions in lesser-developed countries are a result of a conglomeration of policies that are difficult, if not impossible, to derive.

Despite its limited quantitative use, it is worthwhile to use the available survey data to motivate the use of size-based distortions. For instance, Table 1 illustrates how both the incidence of bribes and the percentage of sales paid in bribes initially increases with firm size, but eventually decreases for larger firms in the formal sector. Informal sector firms tell a slightly different story, as bribes increase with firm size in the informal sector. Informal firms pay bribes less often, but when they do, they pay a larger portion of sales. Note that the shape of bribery rates in Table 1 is largely consistent with the pattern described by Tybout (2000). Bribes as a percentage of sales tend to be highest for medium-sized firms.

---

6 The World Bank conducts surveys for both formal and informal firms; however, splicing the surveys together into a coherent picture of the universe of firms remains a daunting project.

7 For example, in Table 2 there is no objective mechanism to ensure that a firm that sees regulation as a “2” represents the same degree of constraint as another firm that rates regulation as a “2.” The two firms could face drastically different regulatory environments, but their baseline for regulation is different.
Table 1: Bribes by Firm Size

<table>
<thead>
<tr>
<th>Sector</th>
<th>Firm Size</th>
<th>Firms Reporting Bribes (%)</th>
<th>As a Share of Sales (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal Sector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Micro (≤10 employees)</td>
<td>49.9</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>Small (10-19)</td>
<td>56.7</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>Medium (20-49)</td>
<td>57.6</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Large (50-249)</td>
<td>58.5</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>Very Large (250+)</td>
<td>55.7</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Informal Sector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small (≤10 employees)</td>
<td>25.5</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Large (10+)</td>
<td>49.1</td>
<td>9.3</td>
<td></td>
</tr>
</tbody>
</table>


Data from the World Bank’s Enterprise Surveys also appears roughly consistent with macroeconomic distortions that vary by firm size. Table 2 shows the mean survey response for each type of firm. Generally it appears that firms are more concerned about taxes, regulation, and corruption as they grow larger; however, there is likely an under-sampling of informal firms (small firms), and small firms are less likely to not have an opinion or not know the answer to each of the survey questions. In totality there seems to be sufficient evidence for studying the effects of size-based distortions. Part of the impetus for doing so in this chapter, is to develop a framework that can be calibrated to individual countries whose specific distortions may vary substantially as exhibited in Table 2. Given the difficulty of quantifying "aggregate" distortions, future calibration exercises testing individual distortions, as in Garcia-Santana and Pijoan-Mas (2014), may be more productive.

---

8Non-responses are excluded from the data presented in Table 1.
Table 2: Survey Results for Firms’ Perceived Constraints

<table>
<thead>
<tr>
<th>Country</th>
<th>Firm Size</th>
<th>Taxes</th>
<th>Regulation</th>
<th>Corruption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>Small</td>
<td>2.39</td>
<td>1.71</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>2.46</td>
<td>1.78</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>2.31</td>
<td>1.84</td>
<td>2.54</td>
</tr>
<tr>
<td>Ecuador</td>
<td>Small</td>
<td>1.97</td>
<td>1.73</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>2.02</td>
<td>1.68</td>
<td>2.52</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>2.08</td>
<td>1.67</td>
<td>1.88</td>
</tr>
<tr>
<td>DRC</td>
<td>Small</td>
<td>2.27</td>
<td>1.93</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>2.40</td>
<td>1.88</td>
<td>2.97</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>2.47</td>
<td>2.21</td>
<td>3.21</td>
</tr>
<tr>
<td>Angola</td>
<td>Small</td>
<td>2.20</td>
<td>1.98</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>2.44</td>
<td>1.99</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>2.41</td>
<td>2.03</td>
<td>2.10</td>
</tr>
</tbody>
</table>

Figures above represent the mean firm response to how large an obstacle was posed by taxes, regulation, and corruption. Rankings are on a scale from 0-4 with 0 representing “no obstacle” and 4 representing “very severe obstacle.” Firms are characterized as small if they have fewer than 20 workers, medium if they have 21-100 workers, and large if they have more than 100 workers. Note that the data used to produce these figures does not generally include firms with fewer than 5 workers which may significantly change some of the averages for small firms. The data is from the World Bank’s Enterprise Surveys for Mexico, Ecuador, Democratic Republic of Congo, and Angola in 2010.

Two additional papers warrant greater elaboration. Restuccia and Rogerson (2008) established several precedents in the literature and began the emphasis on evaluating the effect of distortions on production with heterogeneous firms. They created an environment where heterogeneous productive units faced different prices due to policy-induced distortions. Unlike later authors who study specific policies, Restuccia and Rogerson (2008) develop an undistorted economy and then test the
effects of general distortions on productivity and firm-size distributions. Their calibration exercise utilizes U.S. data, establishing the current practice of considering the U.S. economy to be the undistorted case. Ultimately, the authors find that price heterogeneity caused by distortions can decrease output and TFP by 30% to 50%.

Following soon after, Hsieh and Klenow (2009) has become the de facto standard for comparison in the literature. They follow Restuccia and Rogerson (2008) in considering the U.S. economy to be an undistorted economy, but they specifically calculate TFP dispersion for India and China from micro-level data. They then ask the question, what would be the TFP gains if capital and labor in India and China were allocated similarly to the United States? In this sense, rather than distorting an economy, they show the benefits of policies that would remove distortions already in place. Their results show that TFP gains for China are in the range of 30% to 50% and 40% to 60% in India. While my theoretical model borrows significantly from Hsieh and Klenow (2009), the exercise pursued is certainly different. One of the principal shortcomings in the methodology of Hsieh and Klenow (2009) in the current context is their reliance on survey data. Specifically, the authors are forced to neglect the smallest firms in the firm-size distribution. India’s Annual Survey of Industry omits firms with fewer than 50 people, and China’s Annual Survey of Industrial Production only surveys firms with revenue above $600,000 or that are state-owned. This chapter emphasizes the existence (or lack thereof, as the case may be) of relatively small firms that are excluded from Hsieh and Klenow (2009).

The remainder of the chapter proceeds as follows. The next section develops the theoretical model. The following two sections lay out the scope of the numerical

\[9\] The authors do not claim that the U.S. economy is truly undistorted, but rather that their results should be viewed relative to the U.S.

\[10\] Notice how close these estimates are to those of Restuccia and Rogerson (2008).
application and present the results, respectively. I then discusses the key results. The final section provides conclusions and discusses areas for future work.

3 The Model

This section details the theoretical model. Heterogeneous firms make decisions about how much to invest in innovation and how much labor to hire. Firms have perfect knowledge with regard to the tax rates they face and the results of their investment in innovation. The key object of interest is how firms react to higher tax rates that are specifically tied to the size of their workforce. Intuitively, this changes the incentives for firms to hire workers by changing the marginal profits of each worker, depending on the exact specification for taxes.

3.1 The Economy

The aggregate economy is similar to single-sector version of Hsieh and Klenow (2009). Aggregate output is a CES aggregation of $M$ firms producing differentiated products:

$$Y_t = \left( \sum_{i=1}^{M} \frac{y_{it}}{\sigma} \right)^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1. \tag{1}$$

Output of firm $i$ at time $t$, is $y_{it}$, and $\sigma$ is the elasticity of substitution between intermediate goods. Cost minimization in the production of the final good implies that firms in the intermediate goods sector face a demand curve for their products:

$$p_{it} = \left( \frac{Y_t}{y_{it}} \right)^{\frac{1}{\sigma}}. \tag{2}$$
3.2 Firms

Firms produce differentiated products and compete in monopolistically competitive markets, hence the parameter $\sigma$ for the substitutability of firms’ output in the creation of the final good. Higher levels of $\sigma$ correspond to lower levels of market power for intermediate good producers. Each firm has access to the production technology

$$y_{it} = \varphi_i A_{it} l_{it}^\alpha,$$

where $\varphi$ represents a time invariant productivity draw from a stable distribution as in Melitz (2003). This productivity represents firm specific characteristics like entrepreneurial talent that cannot be improved upon through innovation. Other work in the literature, notably Guner, Ventura, and Yi (2008) and Garcia-Santana and Pijoan-Mas (2014) utilize a Lucas (1978) span-of-control framework to model firm heterogeneity.

I use elements of Melitz (2003) for its simplicity and for comparative purposes with Hsieh and Klenow (2009). Unlike previous works, like Jovanovic (1982) and Hopenhayn (1992), that studied firm heterogeneity, Melitz (2003) abstracts away from the complicated entry and exit process of firms. Specifically, the use of an exogenous death rate of firms greatly simplifies firm optimization while still allowing for heterogeneous production units. Further, since the current chapter seeks to address differences in growth, Melitz (2003) offers an advantage over a Lucas (1978) span-of-control model because there is no decision about how to allocate the population between entrepreneurs and laborers. This eliminates a margin that is not of principal interest here. The framework used here may be of further use in an investigation of the productivity levels required to enter international markets as in Melitz (2003).

11 They also show how their model operates similarly with a Lucas (1978) set-up.
Firm production depends on $A_{it}$, which represents the technology level at firm $i$. This level of technology can be improved over time by innovating (described in Section 3.3). Firms hire labor from an infinite pool of workers who supply their labor inelastically at a fixed wage rate, $w$. The assumption of a fixed wage limits the model to partial equilibrium results; however, Section 6 discusses the implications of endogenizing the wage rate. The exogenous wage rate should be seen as an intermediate step in modeling the aggregate environment, and future work will seek to endogenize it.\(^{12}\)

Timing in the model proceeds as follows (each stage discussed in-depth below). At the beginning of each period, $M$ firms enter the market based on a previous determination that their discounted profit stream from entry exceeded the entry costs.\(^{13}\) Since each firm, $i$, produces a unique good, changes in the number of firms would complicate the model by adding product innovation. Taking aggregate output and the tax rate as given, firms make their production decisions. Aggregate variables are determined by firms’ production decisions, and taxes are collected. At the end of the period, firms face an exogenous probability of death.

### 3.3 Technology and Innovation

Firms are able to improve their technology level, $A_{it}$, by investing in innovation. Innovation follows a process similar to innovation in Atkeson and Burstein (2010); however, in this case there is no uncertainty with respect to the outcomes of innovation.

---

\(^{12}\)The wage rate remains fixed here due to computational complications. Specifically, since each firm faces a different tax rate under size-dependent distortions, aggregate labor demand is the sum of $M$ firms’ non-linear labor demands. Additionally, adding $w$ as an endogenous variable also requires iteration on both $Y_t$ and $w_t$, which slows convergence.

\(^{13}\)The number of firms, $M$, is not time dependent due to the assumptions made with regard to the fixed cost of entry. Fixing the number of firms limits innovation in the model to process innovation.
tion. Firms choose a level of \( q_{it} \in \{0, 1\} \), which determines both the costs and returns to innovation. Firms invest \( A_{it} c(q) \), where \( c(q) \) is the cost function of innovation, which is increasing in the firm’s choice of \( q_{it} \).\(^{14}\) The returns from this investment are a linear combination of a step-size increase and step-size decrease in the level of technology. That is, technology evolves according to

\[
A_{it+1} = q_{it}(A_{it} + \Delta A) + (1 - q_{it})(A_{it} - \Delta A).
\]

This equation can be simplified such that

\[
A_{it+1} = A_{it} + \Delta A(2q_{it} - 1).
\]

This specification for the evolution to technology has several desirable features. First, it captures the fact that firm-level technology can depreciate over time without sufficient levels of investment in innovation. Secondly, the discrete step-size, \( \Delta A \), has clear estimated values in the literature. This allows a key driver of innovation to be closely aligned with the data. Finally, notice that the costs of innovation scale up with the level of technology. This captures the fact that innovation is more costly at higher levels of technology, and it ultimately leads to a steady state in the model. Innovation in the model is best considered process innovation where firms improve their production processes rather than product innovation where firms increase the variety of goods that they produce.\(^{15}\)

\(^{14}\)Note that in steady-state this “choice” is fixed at \( q = \frac{1}{2} \).

\(^{15}\)Process innovation in the model operates in many ways like a Schumpeterian quality ladder model as described in Barro and Sala-i-Martin (2004) in Chapter 7.
3.4 The Firm’s Problem

Firms seek to maximize their discounted stream of profits. Firms discount the future by $\beta (1 - \delta)$, reflecting both a standard discount rate $\beta$ and the probability of firm exit, $\delta$ (explained further in Section 3.5). Each firm, $i$, solves their maximization problem by choosing their price, the amount of labor to hire and how much to invest in innovation according to

$$\max_{q_{it}, l_{it}} \sum_{t=0}^{\infty} (\beta (1 - \delta))^t \left\{ (1 - \tau_{it}) p_{yt} y_{it} - w_{it} l_{it} - A_{it} c(q_{it}) \right\},$$

subject to the evolution of technology in (4), the demand curve for intermediate goods in (2), their production function in (3), and firm characteristics $\varphi_i$ and $A_{i0} > 0$. Distortions in the economy occur through the parameter $\tau$, which represents an implicit tax rate or distortion. These distortions are meant to model the multi-dimensional nature of constraints felt by firms. They represent an aggregation of tax rates, regulation, and corruption. Note that $\tau$ can vary across firms, since the tax rate will later be tied specifically to firm size such that $\tau_{it} = \tau_{it}(l_{it})$.

Firms’ first order conditions are derived for the case where distortions are tied to firm size. In cases where the tax rate is independent of firm size, $\tau'(l_{it}) = 0$. Taking derivatives with respect to the firm’s choice variables, $l_{it}$ and $q_{it}$, results in two first order conditions:

$$w_{it} = (1 - \tau(l_{it})) \kappa Y_{t}^{\frac{1}{\sigma}} (\varphi_i A_{it})^\zeta l_{it}^\xi - \tau'(l_{it}) Y_{t}^{\frac{1}{\sigma}} (\varphi_i A_{it} l_{it}^\alpha)^\zeta,$$  \hspace{1cm} (7a)

$$A_{it} c'(q_{it}) = 2 \Delta A \beta (1 - \delta) \left( (1 - \tau(l_{it+1})) Y_{t+1}^{\frac{1}{\sigma}} (\varphi_i l_{it+1}^\alpha)^\zeta A_{it+1}^{-\frac{1}{\sigma}} - c(q_{t+1}) \right),$$  \hspace{1cm} (7b)

where the scalars $\kappa = \left( \frac{\alpha (\sigma - 1)}{\sigma} \right)$, $\zeta = \frac{\alpha - 1}{\sigma}$, and $\xi = \left( \frac{\alpha (\sigma - 1) - \sigma}{\sigma} \right)$, are used for expositional convenience. Equation (7a) maintains its standard interpretation where the
cost of hiring additional labor, \( w \), is equated to the marginal increase in profits. Note, however, the second term on the right-hand side. This term measures the change in the tax rate faced by a firm that chooses to hire at labor level \( l_{it} \). Equation (7b) equates the marginal cost of innovating today with the discounted marginal profits tomorrow. Since the costs of innovating scale with the technology level, the second term on the right-hand side is necessary because firm investments in innovation today lead to higher costs for innovating in the future. These two first order conditions, along with Equation (5), are sufficient to solve the model in steady-state.

3.5 Entry and Exit

Since I am primarily concerned with the steady-state behavior of firms under different tax distortions, I simplify the entry and exit process of firms in a manner similar to Melitz (2003). Firms face a constant probability of exit, \( \delta \), each period. Firms that die are replaced with identical firms.\(^{16}\) Functionally, the exit process of firms serves to further discount future profits of firms, without adding stochasticity to the steady state. Firm entry typically requires firms to forecast their profits and enter the market if their expected discounted profits exceed the fixed cost of participating in the market. I assume that the \( M \) firms in the quantitative application have already undergone this entry process. Imposing an endogenous entry process adds little to the steady state analysis. Predictably, average productivity in the market will be higher, as the least productive firms have expected discounted profits that are less than the cost of entry.

\(^{16}\)This can be thought of as a perfect markets assumption for firm technology.
3.6 Steady State Equilibrium

**Definition** A steady-state equilibrium in the model is composed of aggregate quantities, \( \{Y^*, T^*\} \), aggregate prices, \( \{w, P\} \), firm-level decisions, \( \{y_i^*, p_i^*, A_i^*, l_i^*, q_i^*\} \), firm-specific tax rates \( \{\tau_i^*\} \), and innovation costs \( \{H, b\} \) such that all firms solve their profit maximization problem and the labor market clears.

Imposing the steady state definition on Equation (5) implies

\[
A^* = A^* + 2\Delta A (2q^* - 1). \tag{8}
\]

Solving for \( q_i^* \) implies that in steady state all firms choose \( q_i^* = q^* = \frac{1}{2} \). Using this condition, and the steady state definition, firms’ first order conditions become

\[
w^* = (1 - \tau(l_i^*))Y^*\frac{1}{2}(\varphi_i A_i^*)^{1/\xi} l_i^{\xi} - \tau'(l_i^*) Y^*\frac{1}{2}(\varphi_i A_i^* l_i^{\alpha})^{1/\zeta}, \tag{9a}
\]

\[
A_i^* c'(q^*) = 2\Delta A (1 - \delta) \left( (1 - \tau(l_i^*)) Y^*\frac{1}{2}(\varphi_i l_i^{\alpha})^{1/\zeta} A_i^* l_i^{-\frac{1}{2}} - c(q^*) \right). \tag{9b}
\]

where \( q^* = \frac{1}{2} \). Given the initial time-invariant productivity draw for the firm, (9a) and (9b) are sufficient to solve for the steady state equilibrium of the economy.

4 Quantitative Application

This section details the parameter values used to numerically solve the model. Each application uses the same parameter values with the exception of the type of implicit tax or distortion that firms face. Note that the purpose of this chapter is not to calibrate results to any particular economy, but to look generally at the effect of different distortions on growth and the innovation process. Further work could evaluate specific distortions and pursue a more detailed calibration exercise. Table 3
serves as a guide to the parameters in the model, and Table 4 provides their specific values used in the calibration.

### Table 3: Parameters of the Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>Elasticity of substitution</td>
<td>$\alpha$</td>
<td>Labor’s share of income</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Firm discount rate</td>
<td>$b$</td>
<td>Innovation costs</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Tax rate on output</td>
<td>$\Delta A$</td>
<td>Innovation step-size</td>
</tr>
<tr>
<td>$M$</td>
<td>Number of firms</td>
<td>$w$</td>
<td>Wage Rate</td>
</tr>
</tbody>
</table>

### Table 4: Parameters Used in Quantitative Application

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>3</td>
<td>$\alpha$</td>
<td>0.33</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.10</td>
<td>$b$</td>
<td>10</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Varies</td>
<td>$\Delta A$</td>
<td>0.1</td>
</tr>
<tr>
<td>$M$</td>
<td>1000</td>
<td>$w$</td>
<td>1</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>$\sim N(10, 5.9234)$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After generating the firm-size distribution for the undistorted economy, I create two tax distortions to evaluate the hypothesis of Tybout (2000). First, I create a linear tax environment where the tax rate is increasing in the amount of labor that firms hire. I set the tax rate so that all firms face a minimum tax rate of 10% (the lowest tax rates observed in the data) and the largest firms in the undistorted economy would face a tax rate of 40% (the approximate mean in the data). This tax scheme runs counter to the intuition of Tybout (2000) regarding the phenomenon

---

17Note that this does not imply the largest firms in the linear distortion environment face a 40% tax rate. Firms in the linear tax environment operate on a smaller scale. If a firm in the linear tax environment decided to utilize the same amount of labor as the largest firm in the undistorted economy, it would face a 40% tax rate.
of the missing middle. This process yields a linear tax schedule of the form \( \tau(l_{it}) = 0.1 + (7.4909 \times 10^{-6})l_{it}. \)

Secondly, I create a quadratic tax schedule to replicate the story of Tybout (2000). The tax rate is increasing for small firms and decreasing for the largest firms. Medium-sized firms face the highest tax rate of 40% (the mean in the data). The smallest firms continue to face a 10% tax rate to capture the lowest rates observed in the data. I fit the curvature of the tax schedule to make sure that no firms receive a subsidy (negative tax rate). The quadratic environment (Tybout-style tax) imposes a tax schedule of the form \( \tau(l_{it}) = 0.1 + (3.1707 \times 10^{-5})l_{it} - (8.3284 \times 10^{-10})l_{it}^2. \)

For reference, Figure 1 plots total tax rates for a number of different developing countries. It illustrates that the tax rates imposed in the quantitative exercise are well within the range of tax rates observed in lesser developed countries. The tax rates, however, include taxes both on profits and on fixed assets. The model imperfectly captures this type of tax by only taxing firms' output. If taxes were to be levied on the technology level of the firm (its only fixed asset), this would likely further decrease TFP as it would increase the expected costs of innovation. Figure 2 plots the exact form of taxes in the linear and quadratic tax environments.

I calibrate the dispersion of time-invariant productivity to approximately match the distribution of educational attainment in China in 2010 for 20-24 year-olds as reported by Barro and Lee (2013). I generate a normal distribution with mean 10 and standard deviation based on the Barro and Lee (2013) data. While this may be a rough proxy for time-invariant productivity or entrepreneurial talent that is not tied to firm technology, it seems more natural than calibrating the distribution of \( \varphi \) to the dispersion of U.S. TFP. That distribution undoubtedly reflects its own long-run adjustment process and may not accurately reflect the distribution of talent in other countries. The number of firms, \( M \), balances the desire for a large economy
with the computational intensity of numerically solving the model. The probability of exogenous exit is set at a standard rate of 10%.

The innovation cost function, \( c(q_{it}) \), is \( bq_{it} \). Importantly, costs are linear in \( q_{it} \), and functional form of \( c(q_{it}) \) is inspired by Atkeson and Burstein (2010). I choose a level of \( b \) such that the economy converges to a steady state equilibrium. To be sure, the level of \( b \) chosen here is not the only level of innovation costs that generates a steady state; however, costs that are too low lead to all firms choosing to innovate with \( q = 1 \), and costs that are too high lead to all firms choosing to pursue no innovation \( (q = 0) \), neither of which implies a balanced growth path. Brandt et al. (2012) estimate firm-level TFP growth in China to be 2.7%. Considering that this is an upper bound for lesser-developed countries, I target \( \Delta A \) more conservatively, so that \( \Delta A \) is approximately 1% of the average level of \( A^* \). Since the equilibrium level of \( A^* \) depends on \( \Delta A \), I iterate over values of \( \Delta A \) until I meet this criteria. Ultimately, this requires that \( \Delta A = 0.1 \).

5 Results

The quantitative application centers on testing five types of distortions. As a baseline, there is an undistorted economy such that \( \tau = 0 \). There are two exercises with tax rates of 20% and 40% across the board (referred to as size-independent taxes). Additionally, there are two exercises with tax rates that depend linearly and non-linearly with firm size (size-dependent taxes). This section begins with size-independent tax environments to establish a baseline and develop intuition and then proceeds to the primary results of interest, tax-dependent distortions.
Source: World Development Indicators 2012. The scatter plot includes 81 randomly chosen countries that have available data on total corporate tax rates and GDP growth for 2012. Countries with total tax rates above 100% are dropped.
Linear and quadratic tax distortions based on firm size. The quadratic environment follows a story similar to Tybout (2000), while the linear environment causes disincentives to hire labor all firm size levels.
5.1 Size-Independent Distortions

Table 5 shows the steady-state levels of aggregate output and aggregate TFP. As expected, the undistorted economy has higher levels of aggregate output and aggregate TFP. The results suggest that imposing a 20% tax rate on a previously undistorted economy would result in a 5.9% loss in log output and a 5.6% loss in log TFP. Similarly, imposing an additional 20% tax on an economy that is in steady state with a 20% tax rate, implies a loss of 8.5% in log output and a 7.5% loss in aggregate TFP.

Table 5: Aggregate Statistics - Size-Independent Distortions

<table>
<thead>
<tr>
<th>Tax Rate</th>
<th>Log Output</th>
<th>Log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undistorted</td>
<td>18.12</td>
<td>11.44</td>
</tr>
<tr>
<td>$\tau = .2$</td>
<td>17.05</td>
<td>10.80</td>
</tr>
<tr>
<td>$\tau = .4$</td>
<td>15.60</td>
<td>9.93</td>
</tr>
</tbody>
</table>

Aggregate TFP is calculated as $\sum_{i=1}^{M} \varphi_i A^*_i$.

Economies with larger distortions have smaller, less technologically advanced firms. Figure 3 illustrates the distribution of firm size across the different size-dependent tax environments, and Table 6 provides statistics on firm characteristics of each environment. The effect of firm size can be seen directly from the firm’s first order conditions. In the size-independent environments, there is no marginal change in the tax rate from hiring more labor; therefore, the second term in Equation (7a) is zero. For a given level of technology and a fixed wage rate, an increase in $\tau$ requires a decrease in the amount of labor that the firm hires. Although size-independent taxes have significant effects on the accumulation of technology in steady state, they generally do not distort the firm size distribution in Figure 3. The firm size distributions tend to simply shift to a smaller scale as the tax rate increases.
Figure 3: Firm Size Distribution across Size-Independent Distortions

Each histogram is created using 50 bins. Note that the x-axis is scaled differently across histograms.
Table 6: Size-Independent Distortions

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Avg. Technology</th>
<th>Avg. Firm TFP</th>
<th>Avg. Firm Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undistorted</td>
<td>9.02</td>
<td>93.20</td>
<td>16,439</td>
</tr>
<tr>
<td>$\tau = .2$</td>
<td>4.72</td>
<td>49.24</td>
<td>4,526</td>
</tr>
<tr>
<td>$\tau = .4$</td>
<td>1.96</td>
<td>20.62</td>
<td>796</td>
</tr>
</tbody>
</table>

5.2 Size-Dependent Distortions

More interesting are the size-dependent distortions. These distortions seek to mimic (quadratic tax) or contrast (linear tax) with the conditions described by Tybout (2000). Specifically, the quadratic distortion is intended to capture the desire for firms to be either small enough to avoid taxes or large enough to influence policy (tax rates). The other size-dependent distortion is linear in firm size. That is, firms face a linearly increasing tax rate as they hire more workers. Figure 2 illustrates the two size-dependent distortions. While these distortions are admittedly somewhat ad hoc, they capture important features to test whether the phenomenon of the missing middle can be captured by size-based tax distortions.

Table 7: Aggregate Statistics - Size-Dependent Distortions

<table>
<thead>
<tr>
<th>Tax Rate</th>
<th>Log Output</th>
<th>Log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undistorted</td>
<td>18.12</td>
<td>11.44</td>
</tr>
<tr>
<td>Linear</td>
<td>17.04</td>
<td>5.06</td>
</tr>
<tr>
<td>Quadratic</td>
<td>17.95</td>
<td>6.90</td>
</tr>
</tbody>
</table>

Aggregate TFP is calculated as $\sum_{i=1}^{M} \varphi_i A_i^*$.  

The striking feature of the size-dependent applications is illustrated in Table 7. It illustrates that the size-dependent distortions have much smaller output losses than the size-independent distortions. The imposition of the linear tax rate results in a
6.0% loss in log output, which is very similar to the size-independent distortion of \( \tau = .2 \). Even more surprising, the quadratic environment only entails a .9% loss in log output. Importantly, however, notice that there are losses in terms of log TFP. The linear distortion implies a 55.8% loss in aggregate TFP, and the quadratic distortion implies a 39.7% loss in TFP.

Table 8 highlights the key differences among firms in the size-dependent distortion economies. The linear economy behaves much the same as the size-independent environment with a tax rate of 20%. Output, technology, and firm size are all significantly lower than in the undistorted economy. However, there is an interesting trade-off between technology and labor in the quadratic environment. Firms in the quadratic tax environment are, on average, larger in steady-state than their counterparts in the undistorted economy. This results because the demand for labor is augmented by tax incentives (i.e. less effective tax) for increasing firm size. This comes at the expense of technology accumulation. The quadratic tax environment has lower accumulated technology and firm level TFP than the undistorted economy. Additionally, it outperforms both size-independent tax schemes along all the dimensions of interest.

Table 8: Size-Dependent Distortions

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Avg. Technology</th>
<th>Avg. Firm TFP</th>
<th>Avg. Firm Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undistorted</td>
<td>9.02</td>
<td>93.20</td>
<td>16,439</td>
</tr>
<tr>
<td>Linear Tax</td>
<td>4.90</td>
<td>50.63</td>
<td>4,026</td>
</tr>
<tr>
<td>Quadratic Tax</td>
<td>6.71</td>
<td>69.01</td>
<td>25,669</td>
</tr>
</tbody>
</table>

Despite the higher initial tax rates faced by firms in the quadratic tax environment, the economy performs similarly to the undistorted economy in terms of total output. The optimal response of firms in the quadratic tax environment is to be-
come larger firms that incur lower tax rates. The distribution of firm size for the size-dependent distortions is presented in Figure 4. It is striking that the economy has no small firms, which is at odds with observations made in lesser developed countries. A discussion of the dynamics that generate these results is reserved for the next section. A clear implication of Figure 4 is that additional frictions are required to incentivize some firms to stay small. Small firms are a crucial part of the “missing middle” phenomenon. Firms in the quadratic environment operate with much higher levels of labor to take advantage of lower tax rates. With larger firms, and less variance in firm size, the quadratic economy shows little loss in terms of log output despite utilizing far less technology in steady state.

Figure 4: Firm Size Distribution across Size-Dependent Distortions

Each histogram is created using 50 bins. Note that the x-axis is scaled differently across histograms.
Table 9 illustrates an additional facet of the results: the dispersion of productivity. The model generates firm TFP levels, $\varphi_iA_{it}$, that are less dispersed than has been observed in the literature. Specifically, Hsieh and Klenow (2009) report the dispersion of productivity for India and China. Their estimates are likely a lower bound for dispersion since their data does not include small firms. Firms in the quadratic economy show less dispersion in both firm size and TFP. This result is driven by their desire to take advantage of lower taxes on the right-hand side of the quadratic distortion.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Undistorted</th>
<th>Quadratic</th>
<th>India</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>90th to 10th</td>
<td>3.31</td>
<td>2.30</td>
<td>5.0</td>
<td>4.9</td>
</tr>
<tr>
<td>75th to 25th</td>
<td>1.81</td>
<td>1.53</td>
<td>2.2</td>
<td>2.3</td>
</tr>
</tbody>
</table>

90th to 10th is the ratio of firm productivity in the 90th percentile to that of the 10th percentile, among entering firms. Likewise, 75th to 25th is the ratio of productivity between the 75th percentile and 25th percentile. The final two columns present the results for India and China in Hsieh and Klenow (2009) for total factor revenue productivity.

Despite the lack of small firms, and seemingly competitive performance of the quadratic tax environment to the undistorted economy, it is important to realize the implications of the differences in aggregate TFP. As the seminal work by Hall and Jones (1999) documents, it is differences in TFP that drive differences in output per worker and income per worker. Table 10 illustrates that the quadratic economy has comparable output in terms of the final good because it utilizes a larger labor force, which decreases output per worker below that of the undistorted economy.
6 Discussion

The results section of the model underscored the potential effect of taxes, both size-independent and size-dependent, on the accumulation of technology. Despite the fact that total output is similar between the undistorted economy and the quadratic tax environment, the accumulation of technology suffers and firms become dependent on labor. This section explores the principal drivers of these results and explores possible extensions.

The assumption of an infinite labor supply and fixed wage rate are fundamentally important in the comparison between the undistorted economy and quadratic tax economy. Suppose that instead, there was a finite labor supply with a standard labor market-clearing wage rate. As firms grow in the quadratic tax environment and take advantage of lower tax rates, they hire increasing amounts of labor. In a well-functioning labor market, this would drive up the real wage, partially offsetting the desire of firms to become large, and lowering total output.

The current set-up of the model is interesting in terms of cross-country comparisons. It suggests that given a low wage rate (i.e. relatively fixed because the supply of labor is large), firms that face size-dependent tax schedules may choose to increase output through labor rather than invest in innovation. This outcome may be reflective of large firms in developing countries. Larger firms, especially national champions, enjoy lower tax rates and special protections that accompany their

Table 10: Output per Worker

<table>
<thead>
<tr>
<th>Distortion</th>
<th>Output per Worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undistorted</td>
<td>4.50</td>
</tr>
<tr>
<td>Quadratic Tax</td>
<td>2.45</td>
</tr>
</tbody>
</table>

28
over-sized influence in the economy.

The presence of large firms in the quadratic tax environment is in line with the story of Tybout (2000); however, there are no small firms at the lower end of the distribution. This result is driven partially by only analyzing the steady state. Outside of the steady state, firms could be observed in different stages of growth, presumably implying that there would be some small firms. A more complex entry and exit process, or other additional friction, is necessary to populate the lower tail of the firm-size distribution as observed in the data.

Finally, the distribution of time-invariant productivity is important in determining the final distribution of firm sizes and TFP. The choice of Chinese educational attainment is made because it should roughly capture or be correlated with entrepreneurial characteristics that cannot be improved by innovation. The data for 20-24 year-olds in China, however, may not be indicative of other countries. A large percentage of 20-24 year olds in China complete secondary school (71.18%), a feature that is not typical of developing countries. In order to generate more realistic firm-size distributions and firm-level TFP distributions the time-invariant distribution must have more weight toward the lower end of the distribution. Other lesser-developed countries’ educational attainment profiles may better fit this requirement.

7 Conclusion

This chapter represents a first step in explicitly modeling innovation and how it is affected by size-based distortions. In doing so, it links the innovation and technology evolution process central to many growth models with the expanding literature on size-based distortions in developing countries. Combining heterogeneous production a la Melitz (2003) with innovation is a necessary step to evaluate the changes in the
process of accumulating technology and how it interfaces with aggregate outcomes.

While further research is certainly required, these results suggest tax schemes that take a quadratic form negatively affect the accumulation of TFP, even while maintaining a similar level of aggregate output compared to the undistorted economy. It captures the incentives for firms to become large and take advantage of lower tax rates. Further, it clearly illustrates that the “missing middle” cannot be explained by tax rates alone. In order to generate the “missing middle,” there must be additional incentives or constraints to keep firms small. Additional research should seek to address the absence of small firms. The results show that certain tax policies may drive total output, even while failing to increase TFP, which has been shown to be an important explanation of cross-country income differences.
References


