Distortions, Endogenous Managerial Skills and Productivity Differences

Dhritiman Bhattacharya, Nezih Guner, and Gustavo Ventura†

May 15, 2012

Abstract

We develop a span-of-control model where managerial skills are endogenous and the outcome of investments over the life cycle of managers. We calibrate this model to U.S plant-size data to quantify the effects of idiosyncratic distortions that are correlated with the size of production units, and how these effects are amplified by managerial investments. We find a quantitatively important role for managerial investments. Idiosyncratic distortions that collect about 6% of output reduce steady-state output by 10.6% in our benchmark model, while when skills are exogenous the reduction is only about 4.9%. We also find that the model can account quite well for properties of Japanese size-distribution data, with a model-implied TFP of about 83% of the U.S. Distortions are critical in accounting for the differences in size distribution between the U.S. and Japan.

Key Words: Distortions, Size, Skill Investments, Productivity Differences.

JEL Classification: O40, E23.

†We thank the Associate Editor (Richard Rogerson), two anonymous referees and participants at the mini-conference “Misallocation and Productivity” at the SED meetings in Ghent. We also thank seminar participants at CREI-MOVE Workshop on Misallocation and TFP, Tinbergen Institute and Yale for their comments. Guner acknowledges financial support from Spanish Ministry of Economy and Competitiveness, Grant ECO2011-28822. All errors are ours.
1 Introduction

Why are some countries so much richer than others? From development accounting exercises, the answer depends heavily on Total Factor Productivity (TFP) differences across countries (Klenow and Rodriguez-Clare, 1997; Prescott 1998; Hall and Jones, 1999; Caselli, 2005). Consequently, much work in the last decade or so has been devoted to understanding the determinants of these measured TFP differences. As part of this effort, a growing body of recent literature has studied the consequences of distortions that alter the efficient allocation of resources across production units, and show the extent to which these distortions can have substantial effects on aggregate productivity.\(^1\)

A unifying feature of this literature is the assumption that the distribution of productivity across production units is exogenous, and invariant with respect to distortions. We depart from this practice in this paper. We develop a model where productivity across establishments is *endogenous* and driven by investments in managerial quality. We use the model to quantify the effects of distortions that are correlated with the size of production units, and the extent to which these effects are magnified by managerial investments.

We study a span-of-control model with a life-cycle structure. Every period, a large number of finitely lived agents are born. These agents are heterogeneous in terms of their initial endowment of managerial skills. The objective of each agent is to maximize the lifetime utility from consumption. In the first period of their lives, agents make an irreversible decision to be either workers or managers. If an agent chooses to be a worker, her managerial skills are of no use and she earns the market wage in every period until retirement. If an agent chooses to be a manager, she can use her managerial skills to operate a plant by employing labor and capital to produce output and collect the net proceeds (after paying labor and capital) as managerial income. Moreover, managers invest resources in skill formation, and as a result managerial skills grow over the life cycle. This implies that a manager can grow the size of her production operation and managerial income by investing a part of her current income each period in skill formation. As managers age and accumulate managerial skills, the distribution of skills (and productivity of production units) evolves endogenously, and

could be affected by distortions and features of the environment.

In the model, the evolution of managerial skills and hence plant size depends not only on initially endowed skills, but also on skill investment decisions. These investment decisions reflect the costs (resources that have to be spent rather than being consumed) and the benefits (the future awards associated to being endowed with better managerial skills). A central assumption in our model is that there are complementarities between skills and investments: managers born with high skills find it optimal to invest more in skills over their lifetime than managers born with low skills. This model property amplifies initial heterogeneity in skills, and leads to increasing dispersion with age in the size of production plants that managers can operate.

We subsequently introduce idiosyncratic distortions in the model. At the start of life, all agents draw an output tax that applies if they become managers. They draw this tax from a distribution that is conditioned on their endowment of managerial ability: a higher initial managerial ability implies that a higher distortion is more likely. This results in distortions that are positively correlated with the productivity and size of production units. In this context, idiosyncratic distortions have broadly two effects. First, a standard reallocation effect, as the enactment of distortions implies that capital and labor services flow from distorted to undistorted production units. Second, a novel skill accumulation effect, as distortions affect the patterns of skill accumulation and thus, the overall distribution of managerial ability (plant productivity). We then ask: What are the quantitative implications of these distortions for output, productivity and the size of establishments? What is the role of managerial investments in amplifying the effects of distortions on resource allocation? What is the interplay between distortions and variation in exogenous, common-to-all establishments, productivity for output, productivity and plant size?

We calibrate the model to match macroeconomic statistics as well as cross sectional features of the U.S. plant data. We assume for these purposes that the U.S. economy is relatively free of the distortions that we focus on. We find that the model can capture central features of the U.S. plant size distribution, including the upper and lower tails. This is critical; on the one hand, the upper tail of the size distribution accounts for the bulk of the employment and output in the economy. On the other hand, the lower tail of the size distribution accounts for the bulk of the plants in the economy.

We consider a menu of idiosyncratic distortions and evaluate their effects on output, plant
size and notions of productivity. The range of these distortions is given by the implicit tax collections generated in stationary equilibrium. Introducing distortions that collect 6% of output lead to a reduction in aggregate output of about 10.6% and output per establishment by about 38.3%. The effects on Total Factor Productivity (TFP) amount to 3.7%. Such distortions lead to a sizeable changes in mean plant size and to a drastic drop in the share of employment in large plants (100 workers or more). Mean size drops by 32.8% and the employment share of large plants drops from about 46% of the total in the undistorted benchmark to about 27.3% in the distorted case. We find that the amplifying effects of skill investments is substantial: in the absence of skill investments, the same structure of distortions generates a reduction in output of only 4.9% in the long run – about 46% of the reduction in the benchmark case. Therefore, our findings indicate that assuming an exogenous distribution of productivities might lead to a substantially underestimation of the consequences of distortions that affect the allocation of resources across production units.

Our results also indicate non-trivial effects on the size distribution of establishments stemming from variation in exogenous aggregate productivity (common to all establishments). Unlike a traditional span-of-control model, our model implies that changes in exogenous aggregate productivity, as they affect the returns to investments in managerial skills, have effects on occupational choice. We find that reducing exogenous productivity by 25% (50%) leads to a reduction in mean establishment size of about 14.2% (26.2%), and a reduction in the share of employment at large units; from about 46% of total employment in the benchmark case to 37.9% (28.1%).

We finally use the model to assess the combined effects of distortions and exogenous variation in economy-wide productivity. For these purposes, we focus on the case of Japan, an economy with (i) lower output per worker than the U.S., (ii) well-known distortions that affect the size of production units and (iii), substantial differences in the size-distribution of establishments in relation to the United States. How much do distortions vis-a-vis variation in exogenous productivity contribute to account for the differences between Japan and the U.S.? To answer this question, we force the model economy to reproduce Japanese data on mean size and output per worker via idiosyncratic distortions and exogenous productivity variation, and find that the model can account for properties of the Japanese size distribution very well. Our model implies an endogenous level of TFP for Japan that is about 83% of the U.S. level. We find that variation in exogenous productivity alone accounts for only 50%,
60% and 11% of the differences in output, TFP and mean size, respectively; the reminder is due to distortions and the interaction between distortions and productivity differences.

The paper is organized as follows. Section 2 presents the model and the modeling of idiosyncratic distortions. Section 3 discusses the calibration of the benchmark model. Section 4 presents the findings associated to the introduction of distortions and exogenous changes in economy-wide productivity. Section 5 provides a quantification of the amplifying effects of skill investments and the importance of endogenizing the productivity distribution. Section 6 illustrates the quantitative implications of the model when applied to Japan via variation in economy-wide productivity and idiosyncratic distortions. Finally, section 7 concludes.

2 Model

Consider the following life-cycle version of Lucas (1978) span-of-control model. Each period, an overlapping generation of heterogeneous agents are born into economy that lives for \( J \) periods. The objective of each agent is to maximize the present value of lifetime utility from consumption

\[
\sum_{j=1}^{J} \beta^{j-1} \log(c_j),
\]  

where \( \beta \in (0, 1) \) and \( c_j \) is the consumption of an age-\( j \) agent.

Each agent is born with an initial endowment of managerial ability. We denote managerial ability by \( z \), and assume that initial (age-1) abilities are drawn from an exogenous distribution with \( cdf F(z) \) and density \( f(z) \) on \([0, z_{\text{max}}]\). Until retirement age \( J_R \), each agent is also endowed with one unit of time which she supplies inelastically as a manager or as a worker. In the very first period of their lives, agents must choose either to be workers or managers. This decision is irreversible. A worker inelastically supplies her endowed labor time to earn the market wage every period until retirement. The decision problem of a worker is to choose how much to consume and save every period.

A manager’s problem, however, is more complicated. A manager has access to a technology to produce output, which requires managerial ability in conjunction with capital and labor services. Hence, given factor prices, she decides how much labor and capital to employ every period. In addition, in every period, a manager decides how much of his net income to
allocate towards current consumption, savings and investments in improving her/his managerial skills.

We assume that each cohort is $1 + n$ bigger than the previous one. These demographic patterns are stationary so that age-$j$ agents are a fraction $\mu_j$ of the population at any point in time. The weights are normalized to add up to one, and obey the recursion, $\mu_{j+1} = \mu_j/(1+n)$.

**Technology** Each manager has access to a span-of-control technology. A plant comprises of a manager with ability $z$ along with labor and capital,

$$y = Az^{1-\gamma} (k^\alpha n^{1-\alpha})^\gamma,$$

where $\gamma$ is the span-of-control parameter and $\alpha \gamma$ is the share of capital. The term $A$ is productivity term that is common to all establishments.\(^2\) Every manager can enhance her future skills by investing current income in skill accumulation. The law of motion for managerial skills is given by

$$z' = z + g(z, x) = z + z^{\theta_1} x^{\theta_2},$$

where $z'$ is next period’s ability and $x$ denotes investment in skill accumulation. The skill accumulation technology described above satisfies three important properties, of which the first two follow from the functional form and the last one is an assumption. First, the technology shows complementarities between current ability and investments in next period’s ability; i.e. $g_{zx} > 0$. Second, $g(z, 0) = 0$. That is, investments are essential to increase the stock of managerial skills. Finally, there are diminishing returns to skill investments: $g_{xx} < 0$. This naturally requires $\theta_2 < 1$.

### 2.1 Decisions

Consider a stationary environment with constant factor prices $R$ and $w$. Let $a$ denote assets that pay the risk-free rate of return $r = R - \delta$.

**Managers** The problem of a manager of age $j$ is given by

\(^2\)In referring to production units, we use the terms *establishment* and *plant* interchangeably.
\[ V_j(z,a) = \max_{x,a'} \{ \log(c) + \beta V_{j+1}(z',a') \} \]

subject to
\[ c + x + a' = \pi(z,r,w,A) + (1+r)a \quad \forall 1 \leq j < J_R - 1, \]
\[ c + a' = (1+r)a \quad \forall j \geq J_R, \]
and
\[ z' = z + g(z,x) \quad \forall j < J_R - 1, \]

with
\[ V_{J+1}(z,a) = \begin{cases} 0 & \text{if } a \geq 0 \\ -\infty & \text{otherwise} \end{cases}. \]

Given her state \((z,a)\), a \(j\)-years old manager decides how much to save, \(a'\), and how much to invest to enhance her skills. Up to the retirement age \(J_R\), a manager’s income consists of her managerial profits and her assets, while after age \(J_R\) her only source of income is from her assets. We assume that agents (managers as well as workers) can lend or borrow at the interest rate \(r\) as long as they do not die in debt.

Since there are no borrowing constraints, factor demands and per-period profits of a manager only depend on her ability \(z\). Managerial income for a manager with ability \(z\) is given by
\[ \pi(z,r,w,A) \equiv \max_{n,k} \{ Az^{1-\gamma} (k^\alpha n^{1-\alpha})^\gamma - wn - (r + \delta)k \}. \]

Factor demands are given by
\[ k(z,r,w,A) = (A(1-\alpha)\gamma)^{\frac{1}{1-\gamma}} \left( \frac{\alpha}{1-\alpha} \right)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}} \left( \frac{1}{r+\delta} \right)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}} \left( \frac{1}{w} \right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} z, \]
and
\[ n(z,r,w,A) = (A(1-\alpha)\gamma)^{\frac{1}{1-\gamma}} \left( \frac{\alpha}{1-\alpha} \right)^{\frac{\alpha\gamma}{1-\gamma}} \left( \frac{1}{r+\delta} \right)^{\frac{\alpha\gamma}{1-\gamma}} \left( \frac{1}{w} \right)^{\frac{1-\alpha\gamma}{1-\gamma}} z. \]

Substituting these into the profit function, one can show that profits (managerial income) is given by
\[
\pi(z, r, w, A) = A^{\frac{1}{1-\gamma}} \Omega \left( \frac{1}{r + \delta} \right)^{\frac{\alpha}{1-\gamma}} \left( \frac{1}{w} \right)^{\gamma \frac{(1-\alpha)}{1-\gamma}} \frac{1}{z},
\]

where \( \Omega \) is a constant equal to

\[
\Omega \equiv (1 - \alpha) \frac{\gamma (1-\alpha)}{\alpha (1-\gamma)} (1 - \gamma) \frac{1}{z}.
\]

Note that since profits are linear function of managerial ability, \( z \), the impact of additional skills on profits is independent of \( z \), and a function only of parameters, exogenous productivity and prices.

The solution to the dynamic programming problem of a manager is characterized by two conditions. First, the solution for next-period assets, \( a' \), is characterized by the standard Euler equation for asset accumulation

\[
\frac{1}{c_j} = \beta (1 + r) \frac{1}{c_{j+1}}.
\]

Second, the optimality condition for \( x \) and (10) imply the following no-arbitrage condition for investing in physical capital and skills

\[
(1 + r) = \pi_x (r, w, A) g_x(z_j, x_j).
\]

The left-hand side of the above equation is next period’s gain in income from one unit of current savings. The manager can also use this one unit as an investment on her skills. Hence, the term \( g_x(z_j, x_j) \) on the right-hand side stands for the additional skills available next period from an additional unit of investment in the current period. The term \( \pi_x (r, w, A) \) is the additional profit generated from an additional unit of managerial skills. Therefore, the right-hand side is the gain in utility by the \( j \)-period old manager from investing one unit of the current consumption good in skill accumulation. To get a unique interior optimum \( g_{xx} \) must be negative, as assumed earlier. This implies that the marginal benefit of investing in skill accumulation is monotonically decreasing in the level of skill investment while the marginal cost, given by \( (1 + r) \), is constant.

Figure 1 illustrates the optimal decision for skill investments \( x \) at a given age \( j \). As the figure illustrates, a higher level of current (age \( j \)) managerial ability leads to higher skill investments as the result of complementarities built into the production of new managerial
skills. Since this occurs at all ages, given prices, initial heterogeneity in skills is magnified by investments in skill acquisition.

The manager’s problem generates decision rules for savings $a' = a^m_j(z, a)$, investment in managerial skills, $x = x_j(z, a)$, as well as the associated factor demands given by $k = k(z_j, r, w, A)$ and $n = n(z_j, r, w, A)$.

**Workers** The problem of an age-$j$ worker is simpler and is given by

$$W_j(a) = \max_{a'} \{ \log(c) + \beta W_{j+1}(a') \}$$

subject to

$$c + a' = w + (1 + r)a \quad \forall 1 \leq j < J_R - 1$$

and

$$c + a' = (1 + r)a \quad \forall j \geq J_R,$$

with

$$W_{j+1}(a) = \begin{cases} 0 & \text{if } a \geq 0 \\ -\infty, & \text{otherwise} \end{cases}.$$

Let the associated savings decision of a worker be $a' = a^w_j(a)$. Like managers, workers can borrow and lend without any constraint as long as they do not die with negative assets.

The objective of each agent born every period is to maximize lifetime utility by choosing to be a worker or a manager. Let $z^*$ be the ability level at which a 1-year old agent is indifferent between being a manager and a worker. This threshold level of $z$ is given by (as agents are born with no assets)

$$V_1(z^*, 0) = W_1(0).$$

Given all the assumptions made, $V_1$ is a continuous and a strictly increasing function of $z$. Therefore, (14) has a well-defined solution, $z^*$. Figure 2 depicts the solution.

**2.2 Equilibrium**

As we mentioned above, members of each new generation are endowed with managerial ability levels distributed with cdf $F(z)$ and density $f(z)$ on $[0, z^{max}]$. After the age-1, the distribution of managerial abilities is endogenous since it depends on investment decisions of managers over their life-cycle.
Let managerial abilities take values in set $\mathcal{Z} = [z^*, z]$ with the endogenous upper bound $z$. Similarly, let $\mathcal{A} = [0, A]$ denote the possible asset levels. Let $\psi_j(a, z)$ be the mass of age-$j$ agents with assets $a$ and skill level $z$. Given $\psi_j(a, z)$, let

$$f_j(z) = \int \psi_j(a, z) da,$$

be the skill distribution for age-$j$ agents. Note that $f_1(z) = f(z)$ by construction.

Each period those agents whose ability is above $z^*$ work as managers, whereas the rest are workers. Then, in a stationary equilibrium with given prices, $(r, w)$, labor, capital and goods market must clear. The labor market equilibrium condition can be written as

$$\sum_{j=1}^{JR-1} \mu_j \int_{z^*}^{z} n(z, r, w, A) f_j(z) dz = F(z^*) \sum_{j=1}^{JR-1} \mu_j$$

where $\mu_j$ is the total mass of cohorts of age $j$. The left-hand side is the labor demand from $JR - 1$ different cohorts of managers. A manager with ability level $z$ demands $n(z, r, w, A)$ units of labor and there are $f_j(z)$ of these agents. The right-hand side is the fraction of each cohort employed as workers times the total mass of all non-retired cohorts in the economy.

In the capital market, the demand for savings is not only generated by managers renting physical capital. There is an additional demand for savings from managers borrowing funds from the capital market to invest in skill accumulation. The capital market equilibrium condition can be written as

$$\sum_{j=1}^{JR-1} \mu_j \int_{z^*}^{z} k(z, r, w, A) f_j(z) dz = \sum_{j=1}^{J-1} \mu_j \int_{z^*}^{z} a_j^w(a) \psi_j(z, a) dz da$$

$$+ \sum_{j=1}^{J-1} \mu_j \int_{z^*}^{z} a_j^m(z, a) \psi_j(z, a) dz da$$

$$- \sum_{j=1}^{JR-2} \mu_j \int_{z^*}^{z} x_j(z, a) \psi_j(z, a) dz da$$

The left-hand-side of the equation (16) above is the capital demand from $JR - 1$ different cohorts of managers. The first two terms on the right-hand-side are the supply of savings from $J - 1$ different cohorts of managers and workers. The third term is the demand for skills investments from $JR - 2$ different cohorts of managers (a manager will stop investing in his skills the period right before his retirement).
The goods market equilibrium condition requires that the sum of undepreciated capital stock and aggregate output produced in all plants in the economy is equal to the sum of aggregate consumption and savings across all cohorts, and skill investments by all managers across all cohorts.

2.3 Idiosyncratic Distortions

Consider now the environment in which managers face distortions to operate production plants. We model these distortions as output taxes that are dependent on the initial ability level of the manager. As the size of production that a manager can operate is (strictly) increasing in her ability, on average, relatively larger plants will be more distorted than smaller ones. In this sense, distortions will be correlated with productivity and size.

At age 1, each agent with ability level \( z \) makes a draw of an idiosyncratic tax \( \tau \) from a distribution \( D(\tau) \), with support \([0, \tau_{\text{max}}(z)]\). We assume in particular that \( \tau_{\text{max}}(z) = \kappa \frac{z}{z_{\text{max}}} \), with \( \kappa > 0 \). Once a manager is attached to a particular tax (distortion) \( \tau \), this tax remains constant over his/her life cycle, and he/she obtains managerial rents \( \pi(r, w, \tau, A) \) that obey

\[
\pi(z, r, w, \tau, A) \equiv \max_{n,k} \left\{ (1 - \tau)Az^{1-\gamma} \left( k^\alpha n^{1-\alpha} \right)^\gamma - wn - (r + \delta)k \right\}.
\]

We assume further that \( D(\tau) \) is a uniform distribution. Hence, the mean value of output taxes for a manager of type \( z \) is given by

\[
E(\tau|z) = \frac{\kappa}{2} \frac{z}{z_{\text{max}}},
\]

which is increasing in \( z \). To summarize, with distortions the timing of events is as follows: (i) agents are born with managerial skills \( z \); (ii) a distortion level is drawn from distribution \( D \); (iii) occupation choice and decisions take place.

We note that our formulation delivers plant-specific, correlated distortions in a simple, parsimonious way. If \( z \) is low, \( \kappa \frac{z}{z_{\text{max}}} \) is small (for a given \( \kappa \), and individuals are more likely to draw a low tax (distortion) level. As \( z \) increases and approaches \( z_{\text{max}} \), a manager is more likely to draw high levels of taxes. Note also that as distortions are assigned to individuals at the start of life and managers invest at different rates, our formulation allows for the coexistence of managers of similar managerial skills who face different distortions.

Our formulation implies that production taxes distort the choice of capital and labor hired, and thus reduce optimal size measured in either capital or labor used, but leave the
capital-labor ratio unaltered. Since distortions affect managerial rents, they matter for skill accumulation. The key condition for skill accumulation is now

\[(1 + r) = \pi_z(r, w, \tau, A)g_x(z_j; x_j).\]  

(17)

As managerial rents are decreasing in \(\tau\), the marginal benefit from skill investment declines as \(\tau\) increases. Therefore, all the same, higher levels of distortions lead to lower levels of managerial skills.

### 2.4 Discussion

It is important at this point to comment on different aspects of our model, the underlying assumptions and its implications. We start by commenting on the nature of the distortions that we entertain, as central drivers of changes in productivity distributions and in connection with the type of skill investments that we consider.

There are essentially two different ways to model distortions, either as embodied in firms or establishments, or attached to individuals –managers– as we do in this paper. Both possibilities can be important channels for changes in the distribution of productivity across production units and their size, so our focus on the latter channel is by no means exhaustive. Our focus on distortions as attached to individuals has implications not only for the size of plants and their productivity, but also for the dynamics of managerial compensation as we elaborate later. As we model the acquisition of general managerial skills, our modeling of distortions captures the effects of distortions that affect the acquisition of such skills; our framework cannot capture the effects of distortions that have consequences on investments on firm of plant-specific productivity. Our formulation roughly captures, for instance, the role of political connections (or lack thereof) that are tied to individuals and that are potentially transferable across production units.\(^3\) More broadly, our formulation captures the effects of idiosyncratic distortions, such as size-dependent policies for instance, insofar as they affect investments in managerial skills that are of general use. Hence, our formulation does not

---

\(^3\)There is a large literature that studies the value of political connections for firms. Focusing on Malaysia, Fisman (2001) and Johnson and Mitten (2003) argue that a large part of several firms’ market value comes from their political connections. In a study more directly related to misallocation of resources, Khandelwal, Schott and Wei (2011) show that the way in which the Chinese government allocated export quotas for textiles under the Multi-fiber Arrangement, prior to the entry of China into WTO, implied substantial resource misallocation, as the firms that received the quotas were not necessarily the most productive ones.
capture the consequences of distortions on firm-specific investments in brand development, R&D or in the formation of firm-specific types of manager’s or worker’s skills.

Overall, our modeling of distortions and their effects on managerial skill investments captures a subset of the multiple effects that idiosyncratic distortions can have on productivity distributions. It should be viewed as a preliminary step towards the understanding of the effects of idiosyncratic distortions in the determination of productivity distributions.

Amplification and Heterogeneity  
Our modeling of managerial skills as endogenous and dependent on investments in goods, implies that changes in economy-wide productivity levels – the term $A$ that is common to all establishments – has effects on skill investment decisions and thus, on productivity and the size of establishments. This stands in contrast with a standard span-of control growth model (i.e. Lucas (1978)), where changes in exogenous productivity do not generate size or productivity differences as changing $A$ has no effects on occupational decisions.\textsuperscript{4} The consequences of changing economy-wide productivity productivity, however, are different in the current setup. As incomes change across steady states due to economy-wide productivity changes, skill investments change, which translate into changes in aggregate managerial skills. As we show in section 4, this process also implies a change in the value of becoming a worker relative to a manager at the start of life, leading in turn to a change in the number of managers (establishments).

It is important to emphasize that distortions as modeled here, as well as changes in economy-wide productivity, have consequences on skill investments, size and the overall productivity distribution as goods are required in the technology to augment skills. This implies that the effects of changes in those variables are amplified by skill investments. These results closely connect our paper to the recent development and trade literature that considers amplification effects of productivity differences or distortions due to investments in skills and R&D. Examples of these papers are Manuelli and Seshadri (2010), Erosa, Koreshkova and Restuccia (2010), Cubas, Ravikumar and Ventura (2012), Rubini (2011) and Atkeson and Burstein (2010, 2011), among several others.

Indeed, the amplification effects of skill investments within our model depend critically on two parameters: $\theta_1$ and $\theta_2$. This can easily be seen in a simple version of the model with two periods and distortions. Given prices, the decision rule for managerial skill investments

\textsuperscript{4}This requires a Cobb-Douglas production technology as we assume here.
in the first period amounts to

\[ x^* = \Delta \theta_2^{1-\theta_2} z_1^{\theta_1} (1 - \tau)^{1/(1-\theta_2)}, \]

where \( \Delta \) is a constant depending on prices, exogenous productivity \( A \) and model parameters. This implies that the growth rate of managerial skills is given by

\[ \frac{z_2^*}{z_1} = 1 + \Delta \theta_2^{-\theta_2/\theta_1} \theta_2^{-\theta_2} z_1^{\theta_1 + \theta_2 - 1} (1 - \tau)^{1/(1-\theta_2)} \]

The above expression has several implications. First, \( x^* \to 0 \) if \( \theta_2 \to 0 \). In this limiting case, the distribution of productivities is exogenous (given by the initial distribution) and idiosyncratic distortions have only allocative effects, as they do not change the distribution of managerial ability across managers (i.e., \( z_2 = z_1 \)). Second, \( \theta_1 = 0 \) reflects the absence of complementarities in skill production: conditional on \( \tau \), all individuals invest the same amount regardless of their initial managerial ability. Third, skill investments amplify initial heterogeneity in skills if \( \theta_1 + \theta_2 > 1 \). Thus, \( \theta_1 \) is important as in conjunction with \( \theta_2 \) determine whether higher initial ability individuals are expected to display higher growth rates in skills (and incomes) than low ability ones. If \( \theta_1 + \theta_2 = 1 \), given \( \tau \), the growth rate in ability is the same for agents regardless of their initial productivity.\(^5\) Finally, the previous point implies that given \( \theta_2 \), the parameter \( \theta_1 \) might be key in accounting for the importance of large establishments in the data. Changes in this parameter amplify initial heterogeneity in skills, and thus, help concentrating labor and capital at larger plants. We exploit this feature of the model when we analyze how important the amplification channels are in section 4.

**TFP** Our model has important implications for Total Factor Productivity (TFP) that we explore in subsequent sections. Note that since the production structure of the model allows for aggregation, the model implies an aggregate production technology

\[ Y = \frac{AZ}{\equiv TFP} K^{\nu \gamma} L^{\gamma (1-\nu)}, \]

\(^5\)This property resembles the role of learning ability, and its interplay with initial human capital, in accounting for differential growth rates in earnings and the patterns of earnings dispersion over the life cycle. See Huggett, Ventura and Yaron (2006, 2011) and Guvenen and Kuruscu (2009), among others.
where $K$ is aggregate capital, $L$ is the aggregate fraction of the population engaged in regular work (see equation 15) and the term $Z$ is the level of endogenous managerial skills raised to the power $1 - \gamma$, i.e.,

$$Z \equiv \left( \sum_{j=1}^{J_R} \mu_j \int_{z^*}^{z} z_j \psi_j(a, z) \, da \, dz \right)^{1-\gamma}.$$

Hence, TFP is endogenous and amounts to $AZ$. Both distortions and the level of economy-wide productivity $A$ affect TFP. Distortions potentially change the lowest productivity level, $z^*$, and through their effects on skill investments, the entire distribution of managerial skills. Similarly, changes in the level of exogenous productivity affect TFP in two ways. First, a change in $A$ reduce TFP directly. Second, since managerial skills are endogenous and depend on the level of $A$, changes in the economy-wide productivity level affect the distribution of managerial skills due to the amplification effects that we discuss above and induce further changes in TFP.$^6$

Overall, it is worth emphasizing that the effects of distortions on the model-based notion of TFP depend critically on the endogeneity of managerial skills. As we show later, correlated distortions lead to a reduction in the threshold $z^*$, a resulting increase in the number of establishments and a concomitant increase in aggregate managerial skills for a fixed distribution. Therefore, potential reductions in TFP driven by distortions require that the effects of distortions on managerial investments dominate the effects stemming from the increase in the number of production units.

### 3 Parameter Values

We assume that the U.S. economy to be distortion free and calibrate the benchmark model parameters to match central aggregate and cross sectional features of the U.S. plant data. Before discussing the calibration strategy, it is worthwhile to emphasize important features of the U.S. plant size data collected from the 2004 U.S. Economic Census. The average size of a plant in the U.S. was about 17.9. The distribution of employment across plants is quite skewed. As many as 72.5% of plants in the economy employed less than 10 workers.

$^6$There is a growing literature that emphasizes the importance of managerial inputs for firm’s productivity; see Bloom and Van Reenen (2010). See also Burtein and Monge (2009) for the importance of international reallocation managerial know-how in income and welfare differences across countries.
but accounted for only 15% of the total employment. On the other hand, less than 2.7% of plants employed more than 100 employees but accounted for about 46% of total employment. These are key features of the data for our analysis of distortions that are correlated with the size of production units.

We assume that the exogenous skill distribution of newborn agents, \( z_1 \), follows a log normal distribution. Specifically, we assume that \( \log(z_1) \) is normally distributed with parameters \( \mu_z \) and \( \sigma_z \). We let the model period correspond to 10 years. Each cohort of agents enters the model at age 20 and live until they are 80 years old. Agents retire at age 60. Hence, in the model agents live for 6 model periods; 4 as workers or managers and 2 as retirees.

There is a total of 9 parameters to calibrate, as listed in Table 1. The product of two of these parameters, importance of capital (\( \alpha \)) and returns to scale (\( \gamma \)), determines the share of capital in output. We determine the values of capital share in output and the depreciation rate from the data. A measure of capital consistent with the current model on business plants should include capital accounted for by the business sector. Similarly a measure of output consistent with our definition of capital should only include output accounted for by the business sector. The measure of capital and output discussed in Guner et al (2008) is consistent with the current plant size distribution model. Hence we use the value of capital output ratio and the capital share reported in that paper. These values are 2.325 (at the annual level) and 0.317, respectively, with a corresponding investment to output ratio of about 0.178. We choose the population growth rate in the model such that the annual population growth rate is 1.1%. Given a capital output ratio and an investment ratio, our (stationary) law of motion of capital implies a depreciation rate of about 6.7% at the annual level.

After calibrating the depreciation rate and the population growth rate, we have 7 more parameters to calibrate: importance of capital, the parameter governing returns to scale, the discount factor, two parameters of the skill accumulation technology and the mean and variance of the skill distribution. Note that the capital share in the model is given by \( \gamma \alpha \), and since this value has to be equal to 0.317, a calibrated value for \( \gamma \) determines \( \alpha \) as well. Hence we have indeed 6 parameters to determine: \( \gamma, \beta, \theta_1, \theta_2, \mu_z \) and \( \sigma_z \). The resulting parameter values are displayed in Table 1.

At the aggregate level, we want the benchmark model to replicate the capital output ratio in the U.S. economy. At the cross sectional level, the model implied distribution of plants should capture some of the important features of the U.S. plant size distribution discussed
in the beginning of this section. We normalize the mean of the skill distribution to zero and jointly calibrate the 5 remaining parameters to match the following 5 moments of the U.S. plant size distribution: mean plant size, fraction of plants with less than 10 workers, fraction of plants with more than 100 workers, fraction of the labor force employed in plants with 100 or more employees, and the aggregate capital output ratio.\footnote{We approximate the initial distribution of managerial skills using gridpoints that range from $-3\sigma_z$ and $3\sigma_z$.} These moments together with their model counterparts are given in Table 2.

The benchmark model is successful in replicating multiple features of the U.S. plant size distribution. The coefficient of variation of the plant size distribution implied by the skill accumulation model is 3.32 which is close to the corresponding value (3.98) in the data. Indeed, the model is able to replicate properties of the entire plant size distribution fairly well as illustrated in Figures 3 and 4. The success of the model in accounting for the tail of the plant-size distribution is important; as we argued earlier, the bulk of employment is there.

**Skill Investments** In our calibration, the fraction of resources that are invested in skill accumulation is of about 2.5% of GDP in the benchmark economy. Viewed as an intangible investment, this is a relatively small fraction of available estimates for these type of investments. McGrattan and Prescott (2010) calculations, for instance, yield an investment rate in a broad notion of intangibles of about 10.8% of output.

Despite the relatively small fraction of resources devoted to the improvement of managerial skills, the incomes of managers grow significantly with age. In the model economy, a manager who is in his 40s (age 3) earns about 2.7 times as much as a manager who is in his 20s (age 1). How does this compare to U.S. data? To answer this question, we first have to take a stand on who is a manager and who is a worker in the data. On one extreme, one can consider 'chief executives', which amount to about 0.9% of the labor force in 2000. A more comprehensive definition can include all those individuals who are categorized in executive, administrative or managerial occupations.\footnote{We used OCC 1990 classification with occupation codes 4 to 22 counted as managers in the U.S. Census. The set of occupations that are classified as managers include Chief executives and public administrators, Financial Managers, Human resources and labor relations managers, Managers and specialists in marketing and advertising and public relations, Managers in education and related fields, Managers of medicine and health occupations, Postmasters.} This group amounts to about 9.2% of the labor force.
force in 2000. About 5.3% of the workforce are managers in the model, which is right in the middle of these two estimates.

In order to calculate the growth rate of managerial incomes over the life-cycle, we use the broader definition of managers above. In an attempt to control for cohort effects, we use the 1980 U.S. Census to construct real managerial incomes for ages 20-29, the 1990 Census for ages 30-39, and the 2000 Census for ages 40-49. The income measure is the total personal income deflated by the Consumer Price Index (CPI). This procedure reveals that managers’ incomes indeed grow significantly with age, as they grow by a factor of about 2.8, from ages 20-29 to ages 40-49. Alternatively, if we normalize managerial incomes by the aggregate level of labor income in each year in Census data, we obtain a growth factor of about 2.2. If instead we control for potential factors affecting all managers as a group, by normalizing managerial incomes by the overall level of managerial incomes in a given year, we obtain a growth factor is of about 2.1. Our model, estimated exclusively with plant-level data, produces an estimate in the range implied by the Census income data.\(^9\)

4 Results

In this section, we present and discuss the central quantitative findings of the paper. We first introduce distortions as described in section 2.3. Subsequently, we explore the implied responses of our model economy to variations in economy-wide productivity.

4.1 Effects of Idiosyncratic Distortions

We evaluate the effects of idiosyncratic distortions by changing the parameter \(\kappa\), which governs the distribution of distortions. We vary \(\kappa\) across steady states so that implicit tax collections out of idiosyncratic distortions achieve certain values. Specifically, we consider variation in \(\kappa\) from zero to the value that generates collections equal to 7\% of output.\(^\text{10}\)

\(^9\)The data also shows that size grows over the life-cycle of plants. Hsieh and Klenow (2012), among others, document that U.S. manufacturing plants that are more than 40 years old are about 7 times larger than those that are less than 5 years old. These authors also show that the growth rate is much lower in Mexico, where older plants are about twice as large as the younger ones, and nearly absent in India.

\(^\text{10}\)The corresponding values for \(\kappa\) are 0, 0.0414, 0.083, 0.137, 0.207, 0.331, 0.538 and 1.184.
Table 3 and 4 show the main findings for values of idiosyncratic distortions that lead to tax collections of 3% and 6% of output. Figure 5 shows the implications for the wider range of distortions for output and mean size. As Table 3 demonstrates, distortions that lead to tax collections of 6% of output imply a reduction in aggregate output of about 10%, and to a substantial increase in the number of production establishments of about 44.8%. Mean size falls from the benchmark value of 17.7 employees to about 13.2. As a result of these changes, output per establishment drops by much more than the reduction in aggregate output: 38.3%. This occurs as with the introduction of distortions that are correlated with size, relatively large, distorted establishments reduce their demand for capital and labor services, leading to a reduction in the wage rate across steady states. This prompts the emergence of smaller production units, as individuals with low initial managerial ability become managers.

The effects outlined above are also present in the analysis in Guner et al (2008), in the context of a standard span-of-control model with capital accumulation. In the current context with skill investments, the consequences of distortions that affect the size of production establishments are more severe. Distortions have detrimental consequences on skill investments, and on average, managers with higher initial ability are more severely distorted. This contributes to a substantial decline in average managerial ability, which declines by 41.1% under idiosyncratic distortions capturing 6% of aggregate output. This decline in average managerial ability translates into a reduction of the model-based notion of TFP, of about 3.7%.

Indeed, mean ability of managers declines due to two reasons. The first reason is a standard reallocation effect: distortions lead to a reallocation of resources across establishments that leads to lower demand for labor and lower wages and results in the emergence of smaller establishments. As managers of these establishments have initial skills that are below the cutoff level, they reduce average managerial skills. The second reason, a skill accumulation effect, is due to the forces that we highlight in this paper. Distorted individuals, who tend to be the initially most able ones, reduce their investments in managerial skills, leading in turn to a further reduction in the economy-wide level of managerial ability. Investment in managerial skills declines from 2.5% of output in the benchmark economy to 2.2% (1.9%) when distortions capture 6% (3%) of output. Both forces contribute to the decline in aggregate output, and the degree of reallocation of resources from large establishments to small ones.
Indeed, as Table 3 shows, idiosyncratic distortions lead to a substantial redistribution of production across establishments. Idiosyncratic distortions that capture 6% of output lead to a drop in the share of employment accounted for by large establishments (100 and more workers) from about 45.8% to 27.3% of total employment, and an increase in the share of small ones (less than 10 workers) from 17.8% to 27.9%. Not surprisingly, the enactment of distortions leads to age-income profiles for managers that are flatter than in the benchmark economy. Between age 1 and age 3, the income of a manager grows by a factor of 2.7 in the benchmark economy, while the growth factor is 2.5 (2.2) when distortions capture 6% (3%) of output.

**The Magnitude of Distortions** How large are the implicit tax distortions that lead to the reductions in mean size discussed above? As Table 4 shows, mean tax rates are relatively small, as an average rate of 3.3% is needed to capture 6% of aggregate output. Not surprisingly, median distortion rates are lower than average ones.\(^{11}\) Note, however, that the relatively low values of means masks somewhat the magnitude of distortions faced by some managers. Table 4 indicates that while the average distortion is 3.3% when distortions capture 6% of aggregate output, the distortions faced by managers at the top of the establishment-size distribution are higher, in the order of about 6.2%.

**Revenue Neutrality** The quantitative experiments discussed above imply that the revenue collected from the enactment of distortions is not constant. To what extent our findings are driven by this feature? To answer this question, we introduce a flat-rate subsidy on either labor income or managerial income, applied to all individuals, so that the net amount of resources effectively extracted from the economy amounts to zero.

The bottom panel of Table 3 presents the findings for the case of output, TFP and mean size.\(^{12}\) The results clearly show that the reported effects do not crucially depend on whether implicit tax collections are returned to individuals in a proportional way. The distorting effects from the idiosyncratic distortions, not their income effects, are the dominant force in our results.

\(^{11}\)Mean and median tax rates reported correspond to the average of distortions \(\tau\) on managers.

\(^{12}\)The subsidy rates are \(-4.6\%\) and \(-9.5\%\) for idiosyncratic distortions collecting 3% and 6% of output, respectively.
The Importance of Distortions Correlated with Productivity  Previous work has shown that for distortions to matter for aggregates and productivity, they have to be heavier for more productive establishments than for less productive ones (Guner et al (2008) and Restuccia and Rogerson (2008)). Does this result still hold in the current environment? Note that distortions as considered here, even if the same for all, affect investment in skills and thus may matter.

To assess the importance of the connection between productivity and distortions in the current environment, we conduct the following experiment: we apply the average tax rates that are required to capture 3 and 6 percent of output to all managers. These average tax rates are about 1 and 3.3%, respectively. We find that the reductions in aggregate output are much smaller than in our benchmark, while leaving mean establishment size essentially unchanged. We conclude from this exercise that the positive association between size and distortions is a central force in the current environment as well.

4.2 Variation in Economy-wide Productivity

We now consider the effects of changes in economy-wide productivity levels; the term A that is common to all establishments. We do this for multiple reasons. First, as we discussed earlier, when goods matter for the acquisition of skills, variation in exogenous productivity affects managerial skills and have effects on occupational choice and plant size. Second, there is substantial variation in the size of establishments across countries that is correlated to the level of development. As it is well documented, productivity differences are a central factor in accounting for the large observed disparities in income across countries. While the mean size of establishments is about 17.9 employees in the U.S. and 15.0 in Norway, it amounts to about 9.7 in Japan. The size of typical establishments is substantially lower in poorer countries. For instance, mean establishment size is 7.5 employees in Turkey, 4.3 in Jordan and 4.4 in India.\footnote{See Bhattacharya (2010) for a documentation of differences in establishment size across countries.} To what extent plant size, output and TFP are affected by exogenous aggregate productivity is a question that can be addressed in the current setup.

We consider two exogenous reductions in productivity (A) relative to the benchmark case: 25% and 50%. Results are presented in Table 5. Not surprisingly, reducing productivity by 25% (50%) leads to a substantial reduction in output across steady states of 38.5%
(68.3%). Changes in the level of productivity also affect the distribution of plant size: reducing productivity by 25% (50%) leads to an increase in the number of establishments by about 15.5% (32.8%) with a corresponding decline in the mean size of establishments.

Quantitatively, the size-distribution effects associated to a reduction in productivity are not trivial. As Table 5 demonstrates, a 25% (50%) reduction in productivity reduces mean managerial ability by 28.5% (49.4%). It leads to a reduction in the share of employment accounted for by establishments with 100 workers or more from about 46% to 37.9% (28.1%), and to an increase in the share of establishments of less than 10 workers from 17.8% to 20.9% (24.5%). Nevertheless, as the table demonstrates, the bulk of endogenous TFP changes are driven by the changes in exogenous productivity $A$.

Overall, changes in aggregate productivity concentrate employment at smaller production units as idiosyncratic distortions do. As a result, they can matter in quantitatively accounting for the observed cross-country differences alongside idiosyncratic distortions correlated with plant’s productivity. In Section 6, we explore in detail the interplay between distortions and exogenous productivity differences for the case of Japan.

5 Amplification Effects: The Importance of Skill Investments

In this section we quantify the importance of the mechanism that we highlight in the paper. We ask the extent to which the effects on output, productivity and the size of plants depend on the presence of skill investments by managers.

The importance of skill investments is intimately connected with the magnitude of the parameter $\theta_2$ in the skill accumulation technology. The parameter $\theta_1$ also plays a role, as it controls the extent of complementarities between skills and investments. If $\theta_2 \to 0$, managerial investments approach to zero, managerial skills become exogenous and invariant to changes in the environment. Overall, $\theta_2$ affects directly the amount of resources invested in skill accumulation and as we discussed earlier, in conjunction with $\theta_1$, it affects the interplay between initially-endowed skills and the growth in income and skills over the life cycle of managers.

In order to evaluate the amplifying role of skill investments, we reduce $\theta_2$ so as to reduce the resources invested in skill investments as a fraction of output, from the benchmark value
(2.5%) to 2%, 1% and 0% (the exogenous skills case). We keep all other parameters fixed at their benchmark values.\textsuperscript{14}

We start by noting that if $\theta_2$ is reduced so that resources in skill formation amount to only 1% of output, the effects on plant productivity and corresponding size distribution statistics are not trivial. Mean establishment size drops from about 17.8 employees to about 13. These changes are largely accounted for by the behavior of the upper tail of the distribution of employment: the share of establishments with 100 workers or more drops from the benchmark value of of 45.8% to 23.4%. This is not surprising: changes in $\theta_2$ affect the growth rate of managerial skills as individuals age, and in conjunction with $\theta_1$, the degree to which these growth rates are positively associated with the level of initial skills. Hence, a reduction in $\theta_2$, for a given distribution of initial managerial ability and $\theta_1$, affects disproportionately the capacity of the model to concentrate employment at the top of the size distribution.

How do different $\theta_2$ values alter the effects of distortions? Table 6 shows the effects of distortions on main variables for different values of $\theta_2$. For each value of $\theta_2$ (the benchmark value, and values that result in 2%, 1% and 0% investment in skills, which we label as CASE I, CASE II and CASE III, respectively), we impose distortions that lead to 6% of resources collected from output in the benchmark economy, and report changes from the corresponding undistorted economies. Figure 6 shows the resulting effects on output for a wider range of distortions considered in previous figures. To highlight the changes under lower values of $\theta_2$, in all cases we use the same parameter values and the same level of distortions as in the benchmark results in section 4.

The distortions considered in the previous sections now have substantially smaller effects on the average values of managerial ability and therefore, on output. Under lower investment in managerial skills, output drops by 6.9% in the intermediate case (1% investment in skills) and by about 4.9% in the exogenous skill case. The decline was 10.6% with the benchmark value of $\theta_2$. Hence, the exogenous skills model captures only about 46% of the output changes implied by distortions. For the wider range of distortions, the results displayed in Figure 6 show, not surprisingly, that the fraction of the output drop captured by the exogenous skill model declines as the severity of the distortions increases. The findings imply that for ‘small’ distortions, a model with exogenous managerial skills (plant-level productivities) is

\textsuperscript{14}The alternative values for $\theta_2$ are 0.363, 0.305 and 0.226, respectively.
not a bad approximation; for 'large', more severe, distortions, the approximation can be a bad one.

Table 6 also shows the changes in establishment size driven by the introduction of idiosyncratic distortions, and the amplification effects for these statistics. For instance, the distortions that lead to collections of about 6% in the benchmark results, lead to a reduction in mean size of 18.4% in the absence of skill investments. The decline was 67.2% in the benchmark calibration. Similarly, while the number establishments rises by 44.8% under the benchmark results, it rises by only 20.3% in the case of no investments.

**The role of complementarities**  We now report on the sensitivity of results in regard to the parameter governing the extent of complementarities in skill accumulation between skills and investments, \( \theta_1 \). As this parameter affects primarily the degree of amplification of initial heterogeneity for the evolution of skills over the life cycle, it substantially impacts the share of employment and output accounted for by the largest establishments. Therefore, we consider different cases of \( \theta_1 \) in terms of their impact on the share of employment of establishments with 100 or more workers. Figure 7 shows the effects of distortions for values of \( \theta_1 \) leading to shares of employment of 30%, 35%, 40%, and the benchmark value (46%). As before, in all cases we use the same parameter values and the same distortions as for the benchmark results in section 4.\(^{15}\)

Overall, the results in Figure 7 indicate that the effects of distortions is not as affected by the extent of complementarities as they were by variation in \( \theta_2 \). In the benchmark case, distortions that capture about 6% of GDP reduce output by about 10.6%. Under a value of \( \theta_1 \) leading to 30% of employment at large establishments, the same distortions imply a reduction in output of about 8.50%, or about 80% of the changes found in the benchmark case.

6 Distortions and their Interplay with Productivity Differences

In previous sections, we showed the quantitative implications of the model in terms of output, productivity measures and the size distribution of establishments. In this section, we use data

\(^{15}\)The alternative values of \( \theta_1 \) are 0.885, 0.901 and 0.917, respectively.
from the Japanese economy to evaluate the performance of the model in several dimensions, and provide quantitative estimates of economy-wide productivity differences between Japan and the United States.

Japan is quite a relevant case to consider from the perspective of this paper. First, it is a relatively rich and educated country. Potential differences in the initial distribution of managerial ability between the U.S and Japan are not a-priori a primary concern. Yet, despite Japan’s income level, its level of output per worker is only about 70% of the U.S. level and it has been at that level for roughly twenty years. Second, there are substantial differences in the size distribution of establishments in relation to the United States. Using data from the Japanese Establishment and Enterprise Census, we calculate that mean establishment size in Japan is substantially below the U.S. level: 9.7 versus 17.9 employees. Not surprisingly, production in Japan is effectively concentrated in small units. As we documented earlier, the fraction of small establishments (less than 10 workers) in the U.S. is 72.5%, the fraction of large establishments (100 workers or more) is about 2.6% and large establishments account for about 46.2% of employment. The corresponding values for Japan differ non trivially: 79.1%, 1.0% and 26.0%, respectively. Finally, Japan is a case that fits well with the case of distortions that are correlated with the size of establishments. As documented in Guner, Ventura and Yi (2006, 2008), Japan regulates severely size of the retail sector at the national level, with policies that go back to the pre-World War II era. Several authors (e.g. Lewis (2005)) document pervasive distortions in other sectors of the Japanese economy. Overall, given our analysis in previous sections, our framework provides a natural vehicle to analyze jointly these differences between the U.S. and Japan, and quantify the importance of distortions and their interplay with exogenous productivity differences.

Using the calibrated parameters of the benchmark economy, we proceed to find the levels of economy-wide productivity ($A$) and distortions as modeled earlier, in order to reproduce, in a stationary equilibrium, two targets: (i) the level of output of Japan relative to the United States; (ii) mean size in Japan. We then contrast the model implications for other properties of the size distribution with Japanese data, and quantify the importance of aggregate productivity differences vis-a-vis idiosyncratic distortions.

**Findings** The findings from the experiment are shown in Table 7. As the table demonstrates, the model is successful in generating the quantitative properties of the Japanese
size-distribution data. The model implies a fraction of small (large) establishments of about 81.6% (0.7%) and a share of employment at large establishments of 24.4%. As mentioned earlier, the corresponding values from the data are 79.1% (1.0%) and 26.0%, respectively. We view these findings as important since they illustrate the capacity of the framework to account for size observations; they give us confidence to take the quantitative findings of the model seriously.

The model generates the Japanese facts with about an aggregate, economy-wide productivity level of about 0.91 (benchmark value 1.0), and distortions that amount to an average of 5.2% across establishments (managers).\textsuperscript{16} Table 7 shows that TFP drops substantially more than the reduction in exogenous productivity; TFP drops by about 17% vis-a-vis a reduction in exogenous productivity of 11%. Hence, there is a significant contribution stemming from the endogenous response of managerial skills across steady states that translate into changes in TFP.

How large are contribution of exogenous productivity differences vis-a-vis distortions in order to generate the results in Table 7? The last column in the Table answers this question. Using the previously found values of $A$, we compute the effects of changes in exogenous productivity only. The results reveal that differences in $A$ play a non-trivial role: they capture nearly half of the changes in output, about 60\% of the changes in TFP and about 11\% of the changes in mean size relative to the benchmark economy. On the other hand, the share of employment at large establishments drops only marginally to 43\% due to the reduction in exogenous productivity.

**Amplification** To what extent the findings for Japan are affected by the endogeneity of managerial skills? Using the distortions and the level of exogenous productivity estimated for Japan, we have also calculated their effects for the case of exogenous skills ($\theta_2 = 0$). In line with previous results, we have found substantially smaller effects on output and the size distribution when managerial skills are exogenous. We found changes in output and the number of establishments of about 23.2\% and 34.7\% when skills are exogenous. The corresponding effects under the benchmark parameters values (column 3 in Table 7) amount to 29.6\% and 74.3\%, respectively. Thus, the model with exogenous skills captures about 78\% of the output changes, and about 47\% of the changes in the number of establishments.

\textsuperscript{16}The value of $\kappa$ for the case of Japan is 1.23.
From all these findings, we conclude that distortions as modeled here matter: exogenous differences in aggregate productivity alone cannot account for the differences in output and the size of establishments between Japan and the United States. Indeed, as the results demonstrate, distortions appear to be the main driver in accounting for differences in the size of establishments between the U.S. and Japan, and a critical factor for the resulting TFP differences between the U.S. and Japan.

7 Conclusion

We developed a span-of-control model where managers invest in the quality of their skills, and used it to quantify the significance of idiosyncratic, correlated distortions, their interplay with aggregate productivity levels, and the extent to which these effects are amplified by skill investments. We found that these distortions can lead to substantial effects on output, productivity and the size distribution. These effects are non-trivially magnified by the endogeneity of managerial skills.

A central implication of our findings is that relatively small distortions on average can lead to substantial effects in the long run. As we elaborate in the text, the endogeneity of managerial skills is key for these results; as the importance of managerial investments is reduced, larger distortions are needed to generate given effects on output and productivity. Overall, our results point to importance of considering models where the distribution of productivities is endogenous in future research. We found that when managerial skills are exogenous, our model substantially underestimates the consequences of distortions on output, aggregate productivity and the distribution of establishment size.

An important implication of our analysis is that exogenous variation in economy-wide productivity is not neutral for occupational choice and thus, the size distribution of establishments. This implies that the model is flexible enough to capture and account for the large differences in size across countries by differences in aggregate productivity, distortions and the interplay between them. For the specific case of Japan, we find that when the model is forced to account for the output gap and the large differences in mean size with the U.S., that differences in productivity alone can capture nearly half of the output gap and 11% of the differences in mean size.

We close the paper with two comments. First, we note that we introduce idiosyncratic,
correlated distortions by assigning them to agents at the start of the life cycle. Managers who are distorted cannot avoid the burden imposed on them, even when no agent is forced to operate a distorted establishment if he/she chooses not to. Perhaps a more realistic approach to the problem would involve the random introduction of distortions in each period that are correlated with managerial productivity. Such an approach, however, would have the drawback of introducing uncertainty in individual decisions, making the comparison between different distortion levels somewhat more involved. Nevertheless, we conjecture that our main quantitative findings would remain unaltered under this extension.

The second comment pertains to the use in detail of observations on the life cycle of managers in the parameterization of the model. We are unaware of empirical work that has focused at the life cycle of managers in detail. The discussion in section 3 indicates that our model is broadly consistent with the growth of managerial income over the life cycle, providing support for our parameterization. In any case, a deeper study that connects an occupational choice model with data on multiple aspects of the life cycle of managers seems warranted. We leave this and other extensions for future work.
References


Table 1: Parameter Values (annualized)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Growth Rate ($n$)</td>
<td>0.011</td>
</tr>
<tr>
<td>Depreciation Rate ($\delta$)</td>
<td>0.067</td>
</tr>
<tr>
<td>Importance of Capital ($\nu$)</td>
<td>0.426</td>
</tr>
<tr>
<td>Returns to Scale ($\gamma$)</td>
<td>0.765</td>
</tr>
<tr>
<td>Mean Log-managerial Ability ($\mu_z$)</td>
<td>0</td>
</tr>
<tr>
<td>Dispersion in Log-managerial Ability ($\sigma_z$)</td>
<td>2.285</td>
</tr>
<tr>
<td>Discount Factor ($\beta$)</td>
<td>0.946</td>
</tr>
<tr>
<td>Skill accumulation technology ($\theta_1$)</td>
<td>0.932</td>
</tr>
<tr>
<td>Skill accumulation technology ($\theta_2$)</td>
<td>0.405</td>
</tr>
</tbody>
</table>

Table 2: Empirical Targets: Model and Data

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Size</td>
<td>17.9</td>
<td>17.7</td>
</tr>
<tr>
<td>Capital Output Ratio</td>
<td>2.325</td>
<td>2.332</td>
</tr>
<tr>
<td>Fraction of Small (0-9 workers) establishments</td>
<td>0.725</td>
<td>0.736</td>
</tr>
<tr>
<td>Fraction of Large (100+ workers) establishments</td>
<td>0.026</td>
<td>0.028</td>
</tr>
<tr>
<td>Employment Share of Large establishments</td>
<td>0.462</td>
<td>0.458</td>
</tr>
</tbody>
</table>
Table 3: Effects of Idiosyncratic Distortions

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Benchmark</th>
<th>3% Output Collected</th>
<th>6% Output Collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Output</td>
<td>100.0</td>
<td>96.8</td>
<td>89.4</td>
</tr>
<tr>
<td>TFP</td>
<td>100.0</td>
<td>99.0</td>
<td>96.3</td>
</tr>
<tr>
<td>Number of Establishments</td>
<td>100.0</td>
<td>115.5</td>
<td>144.8</td>
</tr>
<tr>
<td>Mean Size</td>
<td>100.0</td>
<td>85.8</td>
<td>67.2</td>
</tr>
<tr>
<td>Output per Establishment</td>
<td>100.0</td>
<td>83.8</td>
<td>61.7</td>
</tr>
<tr>
<td>Mean Ability</td>
<td>100.0</td>
<td>82.9</td>
<td>59.9</td>
</tr>
<tr>
<td>Investment in Skills (% of GDP)</td>
<td>2.5</td>
<td>2.2</td>
<td>1.9</td>
</tr>
<tr>
<td>Employment Share (Small, %)</td>
<td>17.8</td>
<td>21.0</td>
<td>27.9</td>
</tr>
<tr>
<td>Employment Share (Large, %)</td>
<td>45.8</td>
<td>37.3</td>
<td>27.3</td>
</tr>
</tbody>
</table>

Revenue Neutrality

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Benchmark</th>
<th>3% Output</th>
<th>6% Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>100.0</td>
<td>98.1</td>
<td>91.6</td>
</tr>
<tr>
<td>TFP</td>
<td>100.0</td>
<td>99.6</td>
<td>97.4</td>
</tr>
<tr>
<td>Mean Size</td>
<td>100.0</td>
<td>87.5</td>
<td>69.5</td>
</tr>
</tbody>
</table>

Uniform Distortions

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Benchmark</th>
<th>3% Output</th>
<th>6% Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>100.0</td>
<td>99.3</td>
<td>97.7</td>
</tr>
<tr>
<td>TFP</td>
<td>100.0</td>
<td>99.8</td>
<td>99.4</td>
</tr>
<tr>
<td>Mean Size</td>
<td>100.0</td>
<td>100.0</td>
<td>97.5</td>
</tr>
</tbody>
</table>

Note: Entries show the effects on displayed variables associated to the introduction of idiosyncratic distortions. The distortions are selected in order to generate implicit tax collections equal to 3% and 6% of aggregate output. ‘Small’ stands for establishments with less than 10 workers whereas ‘Large’ stands for establishments with 100 workers or more.

Table 4: Magnitude of Idiosyncratic Distortions (%)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Benchmark</th>
<th>3% Output Collected</th>
<th>6% Output Collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Distortion</td>
<td>0.0</td>
<td>1.0</td>
<td>3.3</td>
</tr>
<tr>
<td>Median Distortion</td>
<td>0.0</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Mean Distortion (100+ workers)</td>
<td>0.0</td>
<td>4.8</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Note: Entries show the effects on displayed variables associated to the introduction of idiosyncratic distortions. The distortions are selected in order to generate implicit tax collections equal to 3% and 6% of aggregate output.
### Table 5: Role of Aggregate Productivity

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Benchmark</th>
<th>$A = 0.75$</th>
<th>$A = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Output</td>
<td>100.0</td>
<td>61.5</td>
<td>31.7</td>
</tr>
<tr>
<td>TFP</td>
<td>100.0</td>
<td>71.7</td>
<td>45.5</td>
</tr>
<tr>
<td>Number of Establishments</td>
<td>100.0</td>
<td>115.5</td>
<td>132.8</td>
</tr>
<tr>
<td>Mean Size</td>
<td>100.0</td>
<td>85.8</td>
<td>73.8</td>
</tr>
<tr>
<td>Output per Establishment</td>
<td>100.0</td>
<td>53.3</td>
<td>23.8</td>
</tr>
<tr>
<td>Mean Ability</td>
<td>100.0</td>
<td>71.6</td>
<td>50.6</td>
</tr>
<tr>
<td>Investment in Skills (% of GDP)</td>
<td>2.5</td>
<td>2.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Employment Share (Small, %)</td>
<td>17.8</td>
<td>20.9</td>
<td>24.5</td>
</tr>
<tr>
<td>Employment Share (Large, %)</td>
<td>45.8</td>
<td>37.9</td>
<td>28.1</td>
</tr>
</tbody>
</table>

**Note:** Entries show the effects on displayed variables associated to exogenous reductions in the level of economy-wide productivity. ’Small’ stands for establishments with less than 10 workers whereas ’Large’ stands for establishments with 100 workers or more.

### Table 6: Importance of Skill Investments

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Distortions (benchmark)</th>
<th>Distortions (CASE I)</th>
<th>Distortions (CASE II)</th>
<th>Distortions (CASE III)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Output</td>
<td>89.4</td>
<td>90.6</td>
<td>93.1</td>
<td>95.1</td>
</tr>
<tr>
<td>TFP</td>
<td>96.3</td>
<td>97.4</td>
<td>99.5</td>
<td>101.0</td>
</tr>
<tr>
<td>Number of Establishments</td>
<td>144.8</td>
<td>139.2</td>
<td>127.7</td>
<td>120.3</td>
</tr>
<tr>
<td>Mean Ability</td>
<td>58.9</td>
<td>64.3</td>
<td>76.6</td>
<td>86.8</td>
</tr>
<tr>
<td>Mean Size</td>
<td>67.2</td>
<td>70.1</td>
<td>76.7</td>
<td>81.6</td>
</tr>
<tr>
<td>Employment Share (Small, %)</td>
<td>156.5</td>
<td>148.4</td>
<td>135.6</td>
<td>126.6</td>
</tr>
<tr>
<td>Employment Share (Large, %)</td>
<td>59.7</td>
<td>54.0</td>
<td>42.3</td>
<td>29.1</td>
</tr>
</tbody>
</table>

**Note:** Entries show the effects on displayed variables associated to the introduction of idiosyncratic distortions in different cases, when distortions capture 6% of output. The last three columns show the effects driven by the same distortions in the benchmark results, but with a value of the elasticity parameter $\theta_2$ set to generate investments in skills equal to 2% (CASE I), 1% (CASE II) and 0% (CASE III) of output in the absence of distortions. CASE III is the exogenous skill case. ’Small’ stands for establishments with less than 10 workers whereas ’Large’ stands for establishments with 100 workers or more.
Table 7: Japan

<table>
<thead>
<tr>
<th></th>
<th>Benchmark (U.S.)</th>
<th>Japan Data</th>
<th>Japan Model</th>
<th>Japan Model (No Distortions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Output</td>
<td>100.0</td>
<td>70.0</td>
<td>70.4</td>
<td>85.2</td>
</tr>
<tr>
<td>Mean Size</td>
<td>17.7</td>
<td>9.7</td>
<td>9.7</td>
<td>16.8</td>
</tr>
<tr>
<td>Exogenous Productivity (A)</td>
<td>100.0</td>
<td>-</td>
<td>91.0</td>
<td>91.0</td>
</tr>
<tr>
<td>TFP</td>
<td>100.0</td>
<td>-</td>
<td>83.0</td>
<td>88.6</td>
</tr>
<tr>
<td>Number of Establishments (%)</td>
<td>100.0</td>
<td>-</td>
<td>174.3</td>
<td>104.9</td>
</tr>
<tr>
<td>Output per Establishment (%)</td>
<td>100.0</td>
<td>-</td>
<td>40.4</td>
<td>81.2</td>
</tr>
<tr>
<td>Mean Ability</td>
<td>100.0</td>
<td>-</td>
<td>38.9</td>
<td>89.6</td>
</tr>
<tr>
<td>Investment in Skills (% GDP)</td>
<td>2.5</td>
<td>-</td>
<td>1.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Mean Distortion (%)</td>
<td>0.0</td>
<td>-</td>
<td>5.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Small Establishments (%)</td>
<td>74.7</td>
<td>79.1</td>
<td>81.6</td>
<td>74.0</td>
</tr>
<tr>
<td>Large Establishments (%)</td>
<td>2.7</td>
<td>1.0</td>
<td>0.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Employment Share (Large, %)</td>
<td>47.3</td>
<td>26.0</td>
<td>24.4</td>
<td>43.0</td>
</tr>
</tbody>
</table>

Note: Entries show the effects on displayed variables when the model is applied to the case of Japan. The fourth column shows the effects when the distortions and exogenous productivity are varied to match Japan’s output relative to the U.S. and Japan’s mean establishment size. The fifth column shows the effects for the exogenous productivity level previously found, but without distortions. ‘Small’ stands for establishments with less than 10 workers whereas ‘Large’ stands for establishments with 100 workers or more.
Figure 3. Size Distribution

The graph shows the size distribution of plants across different size classes.

- **X-axis**: Plant Size Classes
  - 10-19
  - 20-49
  - 50-99
  - 100-499
  - 500+

- **Y-axis**: Fraction of Plants
  - 0%
  - 2%
  - 4%
  - 6%
  - 8%
  - 10%
  - 12%
  - 14%

The data is represented by a dashed line (Data) and the model predictions by a solid line (Model).
Figure 4. Employment Shares

Data = -
Model = -

Plant Size Classes

Employment Share

0% 5% 10% 15% 20% 25% 30%

10-19 20-49 50-99 100-499 500+
Figure 5. The Effect of Distortions on Output and Mean Size

Output, Benchmark normalized to 100

Mean size

Distortions as a Percentage of Output

- output, right scale
- average size, left scale
Figure 6. The Effect of Distortions on Output: The Role of $\theta_2$
Figure 7. The Effect Distortions on Output: The Role of $\theta_1$