Firm Entry and the TFP Paradox of Productivity-Dependent Distortions

Borja Petit, Miguel Almunia, Juan F. Jimeno, and David Lopez-Rodriguez*

March 5, 2026

Abstract

In economies with heterogeneous firms, distortions correlated with firm productivity can raise aggregate total factor productivity (TFP) even as they worsen the allocation of inputs and the productivity distribution. This productivity paradox arises when distortions sufficiently stimulate firm entry: under decreasing returns to scale, expanded entry raises aggregate TFP by increasing the mass of productive firms. We quantitatively illustrate this mechanism using a model calibrated to Spanish firm-level data, considering two canonical firm-dynamics frameworks: free entry à la Hopenhayn (1992) and occupational choice à la Lucas (1978).

Keywords: firm entry, firm dynamics, total factor productivity, misallocation, size-dependent distortions.

JEL codes: E23, L11, D24.

*Petit: CUNEF Universidad (bpetit@cunef.edu); Almunia: CUNEF Universidad & CEPR (miguel.almunia@cunef.edu); Jimeno: Banco de España, Universidad de Alcalá, CEMFI, CEPR & IZA (juan.jimeno@bde.es); López-Rodríguez: Banco de España (david.lopezr@bde.es). Almunia gratefully acknowledges funding from the Ministerio de Ciencia e Innovación under grants PID2019-111616GA-I00 and RYC2021-031858-I. Diana C. León provided excellent research assistance. The opinions and analyses in this paper are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

Distortions such as size-dependent regulations, preferential tax treatment, or credit subsidies are typically considered as wedges that distort firms' input choices. Seminal contributions such as Restuccia and Rogerson (2008) show that they might generate large productivity losses by weakening selection and by reallocating inputs toward less productive firms and away from more productive ones. This view has become very influential in the interpretation of cross-country productivity differences and in the evaluation of industrial and regulatory policies.

We revisit this question and show that the relationship between correlated distortions—distortions that are increasing in firm productivity—and aggregate TFP depends on firm entry. In some cases, a productivity paradox arises: distortions that are correlated with firms' productivity can increase aggregate TFP, even though they worsen the allocation of productive inputs and the distribution of productivity across firms. By increasing the expected value of operating a firm, distortions raise the equilibrium mass of firms and expand the scale of production in a way that offsets—and in some cases dominates—the standard misallocation and selection effects. Thus, the effect of correlated distortions on aggregate TFP is hump-shaped.

We show this "productivity paradox" (i.e., increasing aggregate TFP with correlated distortions) with a parsimonious model under two standard equilibrium definitions: free entry à la Hopenhayn (1992) and occupational choice à la Lucas (1978). In our model, regardless of the equilibrium framework, distortions affect aggregate TFP through three distinct channels: misallocation of inputs, composition of firms, and the mass of active producers. When firm entry is sufficiently responsive to distortions biased towards small firms, the conventional view that distortions always decrease aggregate productivity is no longer valid. Our results point to the central role of firm entry in policy evaluation.

This paper relates to the broad literature on misallocation and firm-level distortions in heterogeneous-producer economies. Classic contributions emphasize that wedges that vary across firms generate aggregate productivity losses by shifting resources toward low-productivity producers and weakening selection, thereby reducing aggregate TFP (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). A large empirical and quantitative literature has documented the importance of these mechanisms across countries and over time

(Bartelsman et al., 2013) as well as in country-specific contexts (Bhattacharya et al., 2013; García-Santana and Pijoan-Mas, 2014; Gourio and Roys, 2014; Garicano et al., 2016). Our analysis complements this work by showing that, once firm entry is endogenous, correlated distortions can also affect aggregate productivity through a scale (entry) margin that is typically absent in static economies or models with a fixed mass of firms. To our knowledge, the only existing paper that identifies the potentially positive effect of a correlated distortion on aggregate productivity through the firm entry channel is Ando (2021). However, that model focuses on the welfare gains from attenuating risks faced by entrepreneurs, while our setting applies more generally to risk-neutral agents.

More broadly, our paper contributes to the literature on industry equilibrium with entry and exit (Hopenhayn, 1992, 2014) and to occupational-choice models in which entrepreneurship is endogenously determined (Lucas, 1978). A central implication of these frameworks is that policies that alter profitability can induce entry responses that reshape aggregate outcomes. Our results connect directly to work on the macroeconomic implications of size-dependent policies (Guner et al., 2008), which emphasizes how such policies interact with firm dynamics through entry and exit. Our main contribution is to highlight a "productivity paradox": productivity-dependent distortions can raise aggregate TFP even as they worsen misallocation and selection, because they may increase the equilibrium mass of producers in economies with decreasing returns at the firm level.

1 The model

We consider an economy with heterogeneous firms, distortions correlated with firm productivity, and endogenous firm entry. The effect of distortions on aggregate TFP depends critically on how they affect firm entry. To capture this margin, we consider two alternative equilibrium conditions determining the mass of active firms: free entry à la Hopenhayn (1992) and occupational choice à la Lucas (1978).

In an economy à la Hopenhayn (1992), the wage rate is pinned down by the free entry condition (zero profits) while the mass of firms clears the labor market. In an economy à la Lucas (1978), the wage rate is such that the labor market clears while the mass of firms is determined by a productivity threshold below which potential entrants prefer to remain as workers rather than running a firm. One critical difference between these two equilibrium solutions is that, while in Hopenhayn (1992) the wage rate responds to changes in the value of the average entrant, in Lucas (1978) it responds to changes in the value of the marginal entrant. This difference may lead to very different wage and firm entry responses to idiosyncratic distortions.

1.1 Model description

The economy is populated by a continuum of firms of mass M that are characterized by a level of productivity $z \in \mathbb{Z}$. Firms operate a decreasing returns to scale technology combining labor and capital to produce a homogeneous good whose price is normalized to one.[1] Output of firm i is given by:

$$y_i = z_i (k_i^{\alpha} n_i^{1-\alpha})^{\gamma} \tag{1}$$

where γ is the span-of-control parameter, n_i denotes the number of workers, and k_i denotes capital. The wage rate is denoted by w and the capital rental rate is r. For simplicity, we consider a small open economy framework where the rental rate of capital is exogenous.

There are two distortions in the economy. First, firms face a tax of rate τ on profits, which does not distort production but distorts firm entry. Second, firms face an idiosyncratic tax/subsidy on output, ϕ_i. This distortion is potentially correlated with firm productivity, so that $\phi_i = \phi(z)$ with $\phi'(z) \geq 0$.

1.2 The problem of an incumbent firm

The pre-tax profit maximization problem of a firm with productivity z_i, is:

$$\max_{k,n} \; \pi_i \;\; \equiv \;\; (1 - \phi_i)\, y_i - r\, k_i - w\, n_i - c_f \,, \tag{2}$$

[1] An alternative model with constant returns to scale at the firm level and monopolistic competition is isomorphic to ours.

where c_f stands for a fixed cost of production. The solution to this problem is given by a pair of input demand functions that take the form:

$$k_i \;=\; k(z_i, \phi_i) \;=\; \big((1-\phi_i)\, z_i\big)^{\frac{1}{1-\gamma}}\, \gamma\, \Omega(r,w)\, \alpha\, r^{-1}\,, \tag{3}$$

$$n_i \;=\; n(z_i, \phi_i) \;=\; \big((1-\phi_i)\, z_i\big)^{\frac{1}{1-\gamma}}\, \gamma\, \Omega(r,w)\, (1-\alpha)\, w^{-1}. \tag{4}$$

where $\Omega(w, r)$ is a constant that depends on input prices. Given the demand functions in (4), firm's output and (pre-tax) profits are:

$$y_i \;=\; y(z_i, \phi_i) \;=\; \big((1-\phi_i)^{\gamma}\, z_i\big)^{\frac{1}{1-\gamma}}\, \Omega(r,w)\,, \tag{5}$$

$$\pi_i \;=\; \pi(z_i, \phi_i) \;=\; \big((1-\phi_i)\, z_i\big)^{\frac{1}{1-\gamma}}\, \Omega(r,w)\, (1-\gamma) - c_f\,. \tag{6}$$

More details on the derivation of these results can be found in Appendix A.

1.3 Aggregation

Aggregate output in this economy is given by:

$$Y \;=\; \int y(z_i, \phi_i) dF(z_i) \;=\; \text{TFP}\,\left(K^{\alpha} N^{1-\alpha}\right)^{\gamma}, \tag{7}$$

where $F(\cdot)$ is the distribution of firms over z, K and N stand for aggregate capital and labor, respectively,

$$K = \int k(z_i, \phi_i) dF(z_i) \quad \text{and} \quad N = \int n(z_i, \phi_i) dF(z_i)\,,$$

and TFP is aggregate productivity in the economy, which is given by:

$$\text{TFP} = M^{1-\gamma} \left[\int (1-\phi_i)^{\frac{\gamma}{1-\gamma}}\, z^{\frac{1}{1-\gamma}}\, d\,\Gamma(z)\right] \left[\int (1-\phi_i)^{\frac{1}{1-\gamma}}\, z^{\frac{1}{1-\gamma}}\, d\,\Gamma(z)\right]^{-\gamma} \tag{8}$$

where Γ is the cumulative density function (cdf) over z, given by $\Gamma = F/M$.

1.4 Equilibrium à la Hopenhayn (1992)

In this economy, firm's productivity has two additive components, $z_i = \theta_i + \tilde{z}_i$. The first component, θ_i, captures firms' (fixed) potential productivity and is drawn upon entry from a Pareto distribution. The second, $\tilde{z}_i$, follows a standard AR(1) process with normally distributed shocks with parameters (ρ_z, σ_z^2), and captures the firm's deviation from its potential.[2]

Firms operate according to the technology described in the previous section. In addition to this, firms decide whether to exit the market at the end of each period. The value of a firm with productivity $z_i = (\theta_i, \tilde{z}_i)$ is:

$$V(z_i, \phi_i) \;=\; (1-\tau)\,\pi(z_i, \phi_i) \;+\; \beta\,(1-\xi)\,\max\left\{\,\mathbb{E}\left(V(z_i', \phi_i)\,|\,z_i\right),\,0\,\right\} \tag{9}$$

where ξ is an (exogenous) exit probability. The max operator in (9) defines an implicit exit rule of the form $e(z_i, \phi_i) = \mathbf{1}\{\mathbb{E}\left(V(z', \phi_i)\,|\,z_i\right) \leq 0\}$.

Each period, a continuum of potential entrants considers entering. Entry is subject to a cost of c_e units of output. After paying this cost, potential entrants draw an initial productivity from a distribution $\Gamma_0(z)$ that embodies the Pareto distribution of θ_i and a normally distributed shock with mean $\mu_{z,0}$ and variance $\sigma_{z,0}^2$. The problem of a potential entrant is:

$$V_e \;=\; \max\left\{-c_e + \int V(z_i, \phi_i)\,d\,\Gamma_0(z_i)\,,\,0\right\}$$

There is a representative household that consumes, saves and supplies labor services so as to maximize a utility function subject to an aggregate resource constraint. In particular:

$$\max_{C,L}\;\log C \;-\; \nu_0\,\frac{L^{1+\nu_1}}{1+\nu_1}, \quad \text{s.t.}\;\; C = w\,L + \Pi + T \tag{10}$$

where Π is the aggregate profits, and T is a lump-sum transfer from the government. The first-order condition for this problem is $-u_L' = w\,u_C'$.

[2] This is equivalent to a model with productivity following an AR(1) with idiosyncratic shifts drawn from a Pareto distribution. Importantly, the presence of θ–i.e. a Pareto component in firm productivity–is required for the model to correctly replicate the right tail of the firm size distribution in the data.

1.4.1 Equilibrium

Given a profit tax rate τ and a distortion $\phi(z)$, an equilibrium in this economy consists of a wage rate w and a set of policy functions $\{n(z,\phi), k(z,\phi), e(z,\phi)\}$ such that:

(i) the policy functions solve the firms' problem;

(ii) the representative household maximizes utility;

(iii) the labor market clears;

(iv) the free-entry condition holds, and

(v) the distribution of firms is stationary.

1.5 Equilibrium à la Lucas (1978)

There is a unit mass of individuals endowed with a managerial ability $z \in \mathbb{Z}$ drawn from a Pareto distribution. At the beginning of every period, once z is drawn, individuals decide whether to run a firm or work as an employee. If they run a firm, they operate according to the technology described in the previous section and collect (after-tax) profits. If they decide to work as an employee, they get a wage w. Overall, their occupational choice solves:

$$E(z_i, \phi_i) \;=\; \max\,\{\, w\,,\, (1-\tau)\,\pi(z_i, \phi_i)\,\}\,.$$

Because managerial ability does not affect workers' income, this occupational choice implicitly defines a productivity threshold $\hat{z}$ below which individuals prefer to be workers. In particular, $\hat{z}$ satisfies $(1-\tau)\,\pi(\hat{z}, \phi) = w$, so that individuals with managerial ability above $\hat{z}$ run a firm. Thus, the mass of firms is $M = 1 - F(\hat{z})$ and aggregate labor supply is $L = F(\hat{z})$, where F is the cdf of z.[3]

[3]Once distortions are introduced, $\pi(z_i, \phi_i)$ is no longer monotonic in z and a region above the threshold $\hat{z}$ such that individuals may prefer to work as an employee might emerge. In this case, we define M as $M = \int \mathbf{1}\{w < (1-\tau)\pi(z_i, \phi_i)\} dF(z_i)$.

1.5.1 Equilibrium

Given a profit tax rate τ and a distortion $\phi(z)$, an equilibrium in this economy consists of a wage rate w, a productivity cutoff $\hat{z}$, and a set of policy functions $\{n(z, \phi), k(z, \phi)\}$ such that:

(i) the policy functions solve firms' problems;

(ii) the cutoff solves the occupational-choice problem, and

(iii) the labor market clears.

2 Calibration

We calibrate the economy without distortions to micro data on non-financial Spanish firms from the Bank of Spain‘s Central de Balances dataset (Banco de España, 2024). The Central de Balances provides harmonized balance-sheet and income-statement information for Spanish non-financial firms, as reported by firms in the annual accounts filed with the Commercial Registry. The Bank of Spain applies consistent accounting and statistical criteria to clean, standardize, and validate the raw information, ensuring comparability across firms and over time. The resulting panel contains detailed measures of assets, liabilities, revenues, costs, profits, employment, and financing flows (see Almunia et al., 2018).

We set two common parameters exogenously: $\alpha = 0.3$, following standard practice, and the corporate tax rate to 30%.[4] Finally, we set the discount factor to $\beta = 1/(1 + r)$, where r is calibrated internally.

2.1 Hopenhayn economy

Some parameters in our first equilibrium specification are set exogenously. The (inverse) Frisch elasticity in the utility function is set to $\nu_1 = 2$. The scale parameter ν_0 in the disutility of labor supply just scales the economy up or down; consequently, we set its value to 1.

[4] The corporate tax rate level does not affect our results significantly. Note that, in the absence of a profit tax, the effect of distortions on firm entry would be larger than the ones reported in the next section.

The Pareto distribution from which the fixed component of productivity is drawn has two parameters: ξ (shape) and $\underline{z}$ (scale), where ξ controls how fast the probability mass falls as we increase θ and $\underline{z}$ sets a minimum value for θ. We calibrate ξ to match the share of total employment in large firms, defined as firms with 50 or more employees. For comparability, we set the parameter $\underline{z}$ to 1, as this parameter is not identified in the Lucas economy. The span-of-control parameter, γ, is calibrated to match the average firm size, while the interest rate is set such that the capital to output ratio in the economy is equal to 3.

The remaining parameters are calibrated to match the exit rate and key features of the employment distribution. In particular, we match the 10th percentile in the employment distribution, the persistence in employment as well as its variance, and the average and P90-P10 difference in employment among entrants.

2.2 Lucas economy

The economy à la Lucas has fewer parameters; we only need to set a value for γ, ξ, $\underline{z}$ and the interest rate, r. In this economy, changes in $\underline{z}$ are fully absorbed by wages so it is not identified in the model. Consequently, we set $\underline{z} = 1$. For the three remaining parameters, we follow the same identification strategy and choose their values so as to match the average firm size, the share of employment in large firms, and the capital to output ratio. Finally, we assume firms in the Lucas economy face no fixed cost, so we set $c_f = 0$.

2.3 Results

The results of our calibration are reported in Table 1. Overall, the model matches the empirical moments closely. Importantly, it also performs well along untargeted dimensions: both economies replicate the empirical firm-size distribution closely. These untargeted moments indicate that the model captures both the dispersion of firm size and the heavy right tail of the distribution, which are crucial for evaluating the aggregate effects of correlated distortions.

3 Experiments

We now assess the impact of correlated distortions on aggregate TFP by comparing the baseline economy with economies in which firms face idiosyncratic distortions. We define distortions in the spirit of Restuccia and Rogerson (2008). We also consider several counterfactual economies to assess how robust our results are to some of our assumptions and to the value of key parameters.

3.1 Definition of distortions

In an influential paper, Restuccia and Rogerson (2008) study the macroeconomic effects of size-dependent distortions. They define distortions as a subsidy to low-productivity firms and a tax on high-productivity ones. Building on this, we define the distortion ϕ_i as:

$$\phi_i \;=\; \phi(z_i) \;=\; \begin{cases} -\hat{\phi} & \text{if } \; z_i < z^*(\hat{\phi}) \\ \hat{\phi} & \text{if } \; z_i \geq z^*(\hat{\phi}) \end{cases} ,$$

where $z^*(\hat{\phi})$ is the productivity cutoff that makes the policy revenue neutral.[5]

3.2 Decomposition

As stated in equation (8), aggregate productivity is given by:

$$\text{TFP} = M^{1-\gamma} \left[\int (1-\phi_i)^{\frac{\gamma}{1-\gamma}} \, z^{\frac{1}{1-\gamma}} \, d\,\Gamma(z) \right] \left[\int (1-\phi_i)^{\frac{1}{1-\gamma}} \, z^{\frac{1}{1-\gamma}} \, d\,\Gamma(z) \right]^{-\gamma}$$

This expression defines aggregate TFP as a function of three equilibrium objects, $\text{TFP}(\phi, \Gamma, M)$: distortions ϕ, the cdf of firm-level productivity Γ, and the mass of active firms M. Distortions affect aggregate productivity through three channels:

1. *Misallocation channel*: distortions can shift resources toward low-productivity firms, reducing allocative efficiency given aggregate inputs.

[5] In Restuccia and Rogerson (2008), $\tilde{z}$ is given by the median productivity among active firms, which makes their experiment non-neutral in terms of revenue.

2. *Selection channel*: distortions can change the productivity distribution of active firms, Γ, by increasing the share of low-z firms.

3. *Scale channel*: distortions can change the equilibrium mass of firms, M, by affecting the value of entry.

Because aggregate TFP is a non-linear function of (ϕ, Γ, M), a simple one-at-a-time counterfactual would miss the interaction effects among these three channels. To obtain an exact additive decomposition, we apply a Shapley-value procedure. Namely, we define each channel's contribution as its average marginal effect across the $3! = 6$ possible orderings in which the three objects are switched from baseline to distorted values. This yields three components, Δ_ϕ, Δ_Γ, and Δ_M, that satisfy $\Delta\text{TFP} = \Delta_\phi + \Delta_\Gamma + \Delta_M$.

3.3 Results

Table 2 and Figure 1 present the results from our experiment under the two equilibrium definitions we consider. Two patterns stand out. First, aggregate TFP responds non-monotonically to the magnitude of the distortions, measured by $\hat{\phi}$. In both economies, small-to-moderate levels of distortions increase aggregate productivity relative to the undistorted benchmark, despite the fact that distortions worsen the allocation of inputs across firms and weaken firm selection. Second, the shape and persistence of these effects differs across equilibrium frameworks: the Hopenhayn economy exhibits a pronounced hump shape, while the Lucas economy displays positive TFP effects over a wider range of distortions before eventually turning negative sharply at sufficiently large subsidies.

To understand why aggregate productivity may rise in the presence of distortions, we decompose the change in TFP into the three components described in Section III.B.

The *misallocation effect*, captured by Δ_ϕ in Table 2 and Figure 1, is strictly negative and grows in magnitude with the distortion intensity across both equilibrium frameworks.[6] This reflects the canonical mechanism emphasized in the misallocation literature: productivity-dependent wedges distort factor demands, reallocating resources toward less productive firms.

[6] Note that, holding fixed aggregate inputs, the allocation of capital and labor across firms is equivalent across the two equilibrium definition.

The impact on the population of firms—*selection* and *scale* channels—is the crucial force that offsets these efficiency losses for moderate distortions.

Selection and scale effects in the Hopenhayn economy

In the Hopenhayn economy, distortions that subsidize low-productivity firms reduce exit pressures and increase the share of low-z producers that stay active in the economy. As a result, some low-z firms that exit in the baseline economy stay active when distortions are introduced. Thus, average firm-level productivity falls damaging aggregate productivity. This effect corresponds to the *selection channel*, and is captured by Δ_Γ.

Simultaneously, the mass of active firms (which adjusts to clear the labor market) increases, raising aggregate TFP, as captured by Δ_M. This is because distortions reduce average firm size more than they reduce aggregate labor supply, leading to a reallocation of labor across a larger number of smaller firms. To see why, we can derive a closed-form expression for the equilibrium mass of firms using the labor market clearing condition, $N = L$, and the optimality condition for household labor supply, $-u_L = wu_C$. In particular:

$$M = \left(\frac{1}{\bar{n}}\right)^{\frac{\nu_1}{1+\nu_1}} \left(\frac{w}{\nu_0\,\bar{c}}\right)^{\frac{1}{1+\nu_1}}$$

where $\bar{n}$ is the average firm size (or aggregate labor demand when $M = 1$) and $\bar{c}$ denote aggregate consumption for a unit mass of firms, which is given by:[7]

$$\bar{c} = \left(\frac{1-\alpha\gamma}{(1-\alpha)\gamma}\right) w\bar{n} - c_f - c_e\bar{e} \tag{11}$$

where $\bar{e}$ is the average exit rate in the economy, which is independent of M.

Because $\bar{c}$ is linear in $\bar{n}$, the equilibrium mass of firms increases whenever distortions reduce average firm size ($\bar{n}$) sufficiently relative to their effect on wages. For moderate distortions, wages increase to keep the expected value of entry equal to zero. In this case,

[7] Aggregate consumption is: $C = wN + \Pi + T$, which is equivalent to $C = Y - rK - c_f M - c_e M_e$. This expression is linear in M as the equilibrium mass of entrants is $M_e = \bar{e}M$, where $\bar{e}$ is the average exit rate which is unrelated to M. From the firms' optimality conditions, we know that $rk_i = \gamma\alpha(1-\phi_i)y_i$ and $wn_i = \gamma(1-\alpha)(1-\phi_i)y_i$. Aggregating over these expressions yields $rK = \gamma\alpha Y$ and $wN = \gamma(1-\alpha)Y$ since the distortion is defined as revenue-neutral. Altogether implies that $\bar{y} - r\bar{k} = w\bar{n}\,(1-\alpha\gamma)\,/\,((1-\alpha)\gamma)$, and therefore $\bar{c}$ is given by equation (11).

the reduction in average firm size dominates the wage increase, leading to an increase in the mass of firms that offsets the negative *misallocation* and *selection* effects, and delivering positive TFP gains from distortions. For larger distortions, the expected value of entry falls—since more firms need to be taxed for the distortion to be revenue neutral—and wages decline accordingly. However, average firm size contracts more strongly than wages, so labor-market clearing still requires an increase in the mass of active firms, though not enough to compensate for the negative impact of misallocation and selection.

Selection and scale effects in the Lucas economy

In the Lucas economy, firm selection takes place ex-ante by sorting individuals according to their managerial ability. When distortions are introduced, the value of running a firm for the marginal manager increases and, as a result, the productivity threshold for being an entrepreneur falls. This implies that some low-z individuals who would not run a firm in the baseline economy now decide to enter the market, reducing the average firm-level z (*selection channel*) and expanding the share of individuals that run a firm (*scale channel*). The net effect of these two opposing effects is positive for all the $\hat{\phi}$ we consider. However, for distortions above 19.1 percentage points (hereafter, p.p.), the net entry effect is not large enough to compensate for the fall in TFP through misallocation, delivering negative TFP effects.

3.4 Robustness

This section evaluates the robustness of our main quantitative findings. We show that the productivity paradox documented above is not specific to a particular equilibrium environment or parameterization. In particular, we verify that the non-monotonic response of aggregate productivity to distortions and the central role of the entry (scale) margin persist under alternative model settings and alternative measures of aggregate TFP. These exercises therefore confirm that our results reflect a general equilibrium mechanism—the interaction between endogenous entry, selection, and misallocation—rather than a knife-edge feature of the benchmark calibration.

To assess how sensitive our results are to the different assumptions and model parameters, we compute two statistics. First, we report the TFP-maximizing distortion, $\hat{\phi}^{\max}$. To do so, we solve the experiment with $\hat{\phi} = \{1, 2, \ldots, 20\}$ and then find $\hat{\phi}^{\max}$ by quadratic interpolation around the value of $\hat{\phi}$ that yields the highest TFP gain. Second, we report the zero-crossing distortion, $\hat{\phi}^{0}$: the value of $\hat{\phi}$ that generates a zero TFP gain, which we find by linear interpolation around the last $\hat{\phi}$ with positive gains and the first $\hat{\phi}$ with a negative impact on TFP. These statistics are reported in Table 3. More results can be found in Figures 3 and 4.

Alternative TFP measures.

The productivity paradox is not an artifact of the particular definition of aggregate TFP in equation (8): similar qualitative patterns arise using alternative productivity measures commonly used in the literature. Following Guner et al. (2008), we report TFP^{G} as the Solow residual from an aggregate constant-returns-to-scale production function with capital share $\gamma\alpha$. We also report TFP^{E}, an econometrician-style measure computed as the Solow residual from simulated data using the labor share implied by the model. Both measures display responses to distortions that are consistent with the baseline TFP measure, although the magnitudes differ across definitions.

No capital in production.

One simplifying assumption of our model is the exogeneity of the interest rate. This means that part of input costs does not react to the magnitude of distortions. To check whether this assumption can generate the patterns described previously, we run our experiment in an economy where labor is the only production input. We recalibrate both economies following the same identification strategy described in Section II, without targeting the capital-to-output ratio. While the magnitude of the effects is smaller in both equilibrium frameworks, the shape and qualitative properties of our results still hold when labor is the only input of production.

Labor supply elasticity in the Hopenhayn economy.

In the Hopenhayn economy, the wage rate responds to changes in the value of entry, while the mass of active firms responds to distortions to clear the labor market. Thus, the labor supply elasticity, captured by $1/\nu_1$, seems critical for the quantitative response of the mass of active firms. To test how our results depend on this elasticity, we run our experiment for different values of ν_1 and find that, as we decrease the elasticity of labor supply, the effects of distortions on aggregate TFP fall and phase out faster. This is because a less elastic labor supply more tightly constrains the response of the mass of firms to distortions.

The role of decreasing returns to scale.

The degree of returns to scale is also a key driver of the quantitative results. To test how these change with the value of γ, we run our experiment shocking γ by $+5\%$ and -5%. The general pattern is common to both economies: the larger γ, the smaller the TFP gains. For instance, a 10% tax/subsidy rate generates a TFP gain of 1.41 p.p. in the Hopenhayn economy, but this effect falls to 1.02 p.p. when the span-of-control parameter is increased by 5%. In the Lucas framework, this effect falls from 0.92 p.p. to 0.31 p.p.

Interestingly, an increase in the degree of returns to scale also reduces the scope for generating positive TFP effects with the distortion. In the baseline experiment, TFP effects are positive for any distortion below 13.6 p.p. in the Hopenhayn equilibrium and below 19.1 in the Lucas equilibrium. When γ is increased by 5%, the subsidy/tax rate above which TFP effects become negative falls to 12.7 p.p. and 12.8 p.p. respectively.

Productivity distribution.

The key mechanism behind our results is the interaction between the *selection* and *scale* effects. These two effects are linked through the distribution of active firms, $F(z)$. We assume a Pareto distribution for the Lucas economy and for the permanent component of productivity in the Hopenhayn economy. Under this Pareto assumption, it is clear that the slope of the pdf around the minimum productivity for an active firm is crucial. If this slope is small, a significant increase in the number of firms would not imply a large fall in the productivity cutoff. Thus, we could have large *scale effects* without a large *selection effect.*

To test how sensitive our results are to this slope, we run our experiment shocking the shape parameter in the Pareto distribution, ξ, by $\pm 10\%$. The results show that, as we increase the value of ξ, TFP gains from distortions grow. This is because a larger value of ξ makes the right tail of the productivity distribution thinner, so that small changes in the minimum productivity threshold translate into larger changes in the mass of firms, amplifying the *scale effect.*

4 Conclusions

The aggregate productivity consequences of size-dependent distortions in economies with heterogeneous firms depend crucially on the response of firm entry. We develop two benchmark environments—the Hopenhayn (1992) economy with free entry and the Lucas (1978) economy with occupational choice—and introduce a revenue-neutral policy that subsidizes low-productivity firms while taxing high-productivity ones.

Our main finding is a “productivity paradox”: productivity-dependent distortions can increase aggregate TFP—even though they worsen static allocative efficiency and weaken firm selection—by triggering an expansion in the equilibrium mass of active firms. The key mechanism is that distortions raise the expected value of entry, inducing an equilibrium increase in the mass of active firms. Because aggregate productivity depends on both the distribution of firm-level productivity and the scale of entry, the positive entry response can offset the negative effects of misallocation and selection in both model environments, at least for moderate distortions.

We formalize this mechanism through a transparent decomposition of aggregate productivity into misallocation, selection, and scale components. The decomposition highlights that, in both models, misallocation and selection effects are unambiguously negative and grow in magnitude as distortions increase. In contrast, the scale component is positive and potentially large, reflecting the expansion of the number of operating firms. The interplay of these forces rationalizes the non-monotonic (hump-shaped) response of aggregate TFP in the free-entry environment and the more persistent gains in the occupational-choice economy.

These results have implications for interpreting empirical measures of misallocation and for evaluating size-dependent regulations. In particular, policies that appear to distort production away from high-productivity firms may nonetheless coincide with higher aggregate TFP when they induce entry and expand production scale. More broadly, the welfare and efficiency consequences of distortions cannot be inferred solely from static allocative considerations; they depend crucially on how distortions affect firm entry, exit, and occupational choice.

Future work could extend the analysis in several directions. One natural step is to discipline the size-dependent distortions using micro-level policy variation and estimate them jointly with firm dynamics. Another is to incorporate additional margins of adjustment—such as innovation, capital accumulation, or financial frictions—that may interact with entry responses and alter the aggregate implications. Finally, connecting the model-implied productivity decomposition to empirical firm panels would help to assess whether the scale mechanism documented here can quantitatively explain observed cross-country and within-country productivity patterns.

References

Almunia, Miguel, David Lopez Rodriguez, and Enrique Moral-Benito, "Evaluating the macro-representativeness of a firm-level database: an application for the Spanish economy," *Banco de Espana Ocassional Paper*, 2018, (1802).

Ando, Sakai, "Size-dependent policies and risky firm creation," *Journal of Public Economics*, 2021, *197*, 104404.

Banco de España, "Microdata on individual enterprises (CBI)," 2024. BELab Data Laboratory. DOI: https://doi.org/10.48719/BELab.CBI9524_01.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 2013, *103* (1), 305–334.

Bhattacharya, Dhritman, Nezih Guner, and Gustavo Ventura, "Distortions, Endogenous Managerial Skills and Productivity Differences," *Review of Economic Dynamics*, 2013, *16* (1), 11–25.

García-Santana, Manuel and Josep Pijoan-Mas, "The reservation laws in India and the misallocation of production factors," *Journal of Monetary Economics*, 2014, *66*, 193–209.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, November 2016, *106* (11), 3439–79.

Gourio, François and Nicolas Roys, "Size-dependent regulations, firm size distribution, and reallocation," *Quantitative Economics*, 2014, *5* (2), 377–416.

Guner, Nezih, Gustavo Ventura, and Yi Xu, "Macroeconomic implications of size-dependent policies," *Review of Economic Dynamics*, 2008, *11* (4), 721–744.

Hopenhayn, Hugo A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

_ , "Firms, Misallocation, and Aggregate Productivity: A Review," *Annual Review of Economics*, 2014, *6* (Volume 6, 2014), 735–770.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Lucas, Robert E., "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 1978, *9* (2), 508–523.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Figures

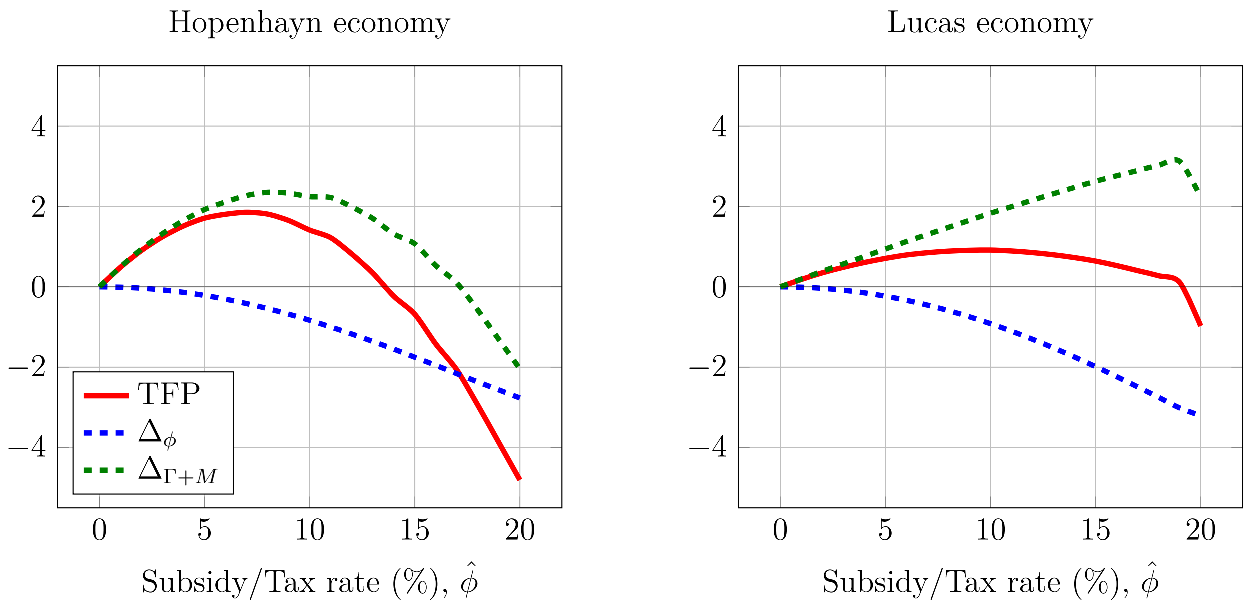

Figure 1: Impact of distortions on aggregate TFP (in p.p. relative to baseline)

Notes: The figure plots the change in aggregate TFP (in percentage points relative to the undistorted baseline) as a function of the subsidy/tax rate $\hat{\phi}$, for $\hat{\phi} \in \{1, 2, \ldots, 20\}$. The distortion subsidizes firms below a productivity cutoff $z^*(\hat{\phi})$ at rate $\hat{\phi}$ and taxes firms above it at the same rate, with $z^*(\hat{\phi})$ chosen to ensure revenue neutrality. TFP is decomposed into three components using a Shapley-value procedure: Δ_ϕ captures the misallocation effect (holding the distribution of firms and their mass fixed), Δ_Γ captures the selection effect (changes in the productivity distribution of active firms), and Δ_M captures the scale effect (changes in the equilibrium mass of firms). The left panel corresponds to the Hopenhayn (1992) economy with free entry; the right panel corresponds to the Lucas (1978) economy with occupational choice. Both economies are calibrated to Spanish firm-level data from the Bank of Spain's Central de Balances.

Tables

Table 1: Model parameters

Parameter		Value	Target	Model	Data
			Hopenhayn (1992) economy		
γ	Firms' returns to scale	0.63	Average firm size	14.4	14.4
ξ	Pareto distribution, shape	2.93	% Employment in large firms	55.7	54.8
r	Interest rate	6.35	Capital to output ratio	3.0	3.0
ρ	Persistence of $\tilde{z}$	0.91	Persistence in employment	0.96	0.96
σ_z^2	Variance of shocks to $\tilde{z}$	0.15	Variance of log employment	1.3	1.2
$\mu_{z,0}$	Average of initial $\tilde{z}$	-0.27	Average employment, entrants	5.0	5.1
$\sigma_{z,0}^2$	Variance of initial $\tilde{z}$	0.01	P90 employment, entrants	7.8	7.2
c_e	Entry cost	10.8	P10 employment, entrants	1.0	1.0
c_f	Fixed cost	0.88	Exit rate	8.0	8.0
		Firm-size distribution:	P10 of employment dist.	0.8	1.0
		(untargeted)	P25 of employment dist.	1.6	2.0
			P50 of employment dist.	3.6	3.6
			P75 of employment dist.	9.2	8.1
			P90 of employment dist.	24.2	19.6
			Lucas (1978) economy		
γ	Firms' returns to scale	0.65	Average firm size	14.3	14.3
ξ	Pareto distribution, shape	2.49	% Employment in large firms	53.6	54.8
r	Interest rate	6.51	Capital to output ratio	3.0	3.0
		Firm-size distribution:	P10 of employment dist.	2.1	1.0
		(untargeted)	P25 of employment dist.	2.6	2.0
			P50 of employment dist.	4.1	3.6
			P75 of employment dist.	9.2	8.1
			P90 of employment dist.	27.4	19.6

Notes: we define large firms as those with at least 50 workers. In the Hopenhayn economy, the fixed cost of 0.88 corresponds to 1.23 times the equilibrium wage, while the entry cost is equivalent to 16 times the equilibrium wage.

Table 2: Impact of distortions on TFP (p.p. relative to baseline)

	Hopenhayn economy					Lucas economy				
$\hat{\phi}$	TFP	Δ_ϕ	$\Delta_{\Gamma+M}$	Δ_Γ	Δ_M	TFP	Δ_ϕ	$\Delta_{\Gamma+M}$	Δ_Γ	Δ_M
2	0.91	-0.03	0.94	-2.66	3.60	0.35	-0.04	0.39	-2.63	3.01
5	1.71	-0.21	1.92	-6.91	8.83	0.71	-0.23	0.94	-6.53	7.47
10	1.41	-0.83	2.24	-14.80	17.05	0.92	-0.92	1.83	-13.24	15.08
15	-0.68	-1.75	1.07	-22.31	23.38	0.64	-1.99	2.63	-20.13	22.76
20	-4.81	-2.77	-2.04	-30.52	28.48	-0.98	-3.21	2.23	-28.06	30.29

Notes: The table reports the change in aggregate TFP and its decomposition (in percentage points relative to the undistorted baseline) for selected values of the subsidy/tax rate $\hat{\phi}$. The distortion subsidizes firms below a revenue-neutral productivity cutoff $z^*(\hat{\phi})$ at rate $\hat{\phi}$ and taxes firms above it at the same rate. TFP is decomposed using a Shapley-value procedure into three additive components: Δ_ϕ (misallocation effect), Δ_Γ (selection effect), and Δ_M (scale effect), which sum to the total change in TFP. $\Delta_{\Gamma+M}$ reports the combined contribution of selection and scale. The left panel corresponds to the Hopenhayn (1992) economy with free entry; the right panel corresponds to the Lucas (1978) economy with occupational choice. Both economies are calibrated to Spanish firm-level data.

Table 3: Sensitivity analysis: zero-crossing and TFP-maximizing $\hat{\phi}$

	Hopenhayn economy			Lucas economy		
	$\hat{\phi}^0$	$\hat{\phi}^{\max}$	ΔTFP	$\hat{\phi}^0$	$\hat{\phi}^{\max}$	ΔTFP
Baseline	13.60	7.02	1.86	19.10	9.71	0.92
TFP measured as TFP^{E}	10.13	5.07	0.71	17.75	8.67	0.71
TFP measured as TFP^{G}	10.13	5.07	1.31	17.75	8.66	1.26
Lower returns to scale ($\gamma \times 0.95$)	14.74	7.24	2.28	19.48	14.02	1.64
Higher returns to scale ($\gamma \times 1.05$)	12.72	6.29	1.52	12.83	6.24	0.46
Fatter-tail z dist. ($\xi \times 0.90$)	12.50	6.27	1.46	13.92	7.06	0.47
Thinner-tail z dist. ($\xi \times 1.10$)	16.10	8.01	2.39	17.55	12.85	1.58
Economy with no capital	13.77	6.73	0.70	15.33	7.61	0.50
Inelastic labor supply ($\nu_1 = 1/0.01$)	12.78	6.66	1.65	–	–	–
Elastic labor supply ($\nu_1 = 1/3$)	15.03	7.53	2.15	–	–	–

Notes: we solve the different economies for $\hat{\phi} = \{1, 2, \ldots, 20\}$. Then, compute the $\hat{\phi}$ at which $\Delta\text{TFP} = 0$ by linear interpolation around the first value of $\hat{\phi}$ with negative TFP effects, and the $\hat{\phi}^{\max}$ (the level of $\hat{\phi}$ that maximizes TFP gains) with a quadratic interpolation around the value of $\hat{\phi}$ that delivers the greatest TFP effect.

SUPPLEMENTAL APPENDIX

Firm Entry and the TFP Paradox of Productivity-Dependent Distortions

Borja Petit, Miguel Almunia, Juan F. Jimeno,
and David Lopez-Rodriguez

March 5, 2026

A Deriving input demands, output and profits

The pre-tax profit maximization problem of a firm with productivity z_i, is:

$$\max_{k,n} \; (1-\phi_i)\, z_i (k_i^{\alpha} n_i^{1-\alpha})^{\gamma} - r\, k_i - w\, n_i \,,$$

where n_i denotes the number of workers, and k_i denotes capital. The first-order conditions for this problem are $\gamma\, \alpha\, (1-\phi_i)\, y_i \;=\; r\, k_i$ and $\gamma\, (1-\alpha)\, (1-\phi_i)\, y_i = w\, n_i$.

A.0.1 Inputs demand

Taking the ratio of the two first-order conditions, delivers the standard capital-to-labor ratio:

$$\frac{k_i}{n_i} = \left(\frac{\alpha}{r}\right) \left(\frac{1-\alpha}{w}\right)^{-1} .$$

Note that, because the distortion is defined as a tax/subsidy to output, it does not affect the capital-labor allocation within the firm. Using this ratio in the first-order condition for labor, yields:

$$\gamma\, (1-\alpha)\, (1-\phi_i)\, z_i n_i^{\gamma} \left(\frac{\alpha}{r}\right)^{\gamma\alpha} \left(\frac{1-\alpha}{w}\right)^{-\gamma\alpha} = w\, n_i \,.$$

Finally, solving for n_i yields:

$$n_i \;=\; \big((1-\phi_i)\, z_i\big)^{\frac{1}{1-\gamma}}\, \gamma\, \Omega(r,w) \left(\frac{1-\alpha}{w}\right) \;=\; n(z_i, \phi_i)\,.$$

where the constant $\Omega(r,w)$ is given by:

$$\Omega(r,w) := \left[\gamma \left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{\gamma}{1-\gamma}} .$$

Similarly, we can derive the optimal capital demand by using $n(z_i, \phi_i)$ in the the optimal capital-to-labor ratio derived before. This yields:

$$k_i \;=\; n_i \left(\frac{\alpha}{r}\right) \left(\frac{1-\alpha}{w}\right)^{-1} \;=\; \big((1-\phi_i)\, z_i\big)^{\frac{1}{1-\gamma}}\, \gamma\, \Omega(r,w) \left(\frac{\alpha}{r}\right) \;=\; k(z_i, \phi_i)$$

A.0.2 Output and profits

Given k_i and n_i, firm's output is:

$$y_i \;=\; z_i \Big(\big((1-\phi_i)\, z_i\big)^{\frac{1}{1-\gamma}}\, \Omega(r,w)\, \gamma \left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{1-\alpha}{w}\right)^{(1-\alpha)} \Big)^{\gamma}$$

which simplifies to:

$$y_i \;=\; \big((1-\phi_i)^{\gamma}\, z_i\big)^{\frac{1}{1-\gamma}}\, \Omega(r,w) \;=\; y(z_i, \phi_i)$$

Finally, using this expression, we can derive pre-tax profits as:

$$\begin{aligned}\pi_i &= y_i - w\,n_i - r\,k_i - c_f \;=\; (1-\gamma)\,(1-\phi_i)\,y_i - c_f \\ &= \big((1-\phi_i)\,z_i\big)^{\frac{1}{1-\gamma}}\,\Omega(r,w)\,(1-\gamma) - c_f \;=\; \pi(z_i,\phi_i)\end{aligned}$$

A.0.3 Aggregate TFP

To derive an analytical expression for aggregate TFP, we start by defining firm i's output as a function of aggregate inputs. Note that:

$$\frac{n_j}{n_i} = \frac{k_j}{k_i} = \frac{\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}}{\big((1-\phi_i)\,z_i\big)^{\frac{1}{1-\gamma}}}$$

and aggregating over k_j and n_j, we get:

$$k_i = \frac{\big((1-\phi_i)\,z_i\big)^{\frac{1}{1-\gamma}}\,K}{M\,\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]}, \qquad n_i = \frac{\big((1-\phi_i)\,z_i\big)^{\frac{1}{1-\gamma}}\,N}{M\,\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]}$$

Using this in the definition of firm's i output gives:

$$y_i \;=\; z_i\left[\frac{\big((1-\phi_i)\,z_i\big)^{\frac{1}{1-\gamma}}\,K}{M\,\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]}\right]^{\alpha\gamma}\left[\frac{\big((1-\phi_i)\,z_i\big)^{\frac{1}{1-\gamma}}\,N}{M\,\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]}\right]^{\gamma(1-\alpha)}$$

so that:

$$y_i \;=\; \frac{z_i\big((1-\phi_i)\,z_i\big)^{\frac{\gamma}{1-\gamma}}}{M^{\gamma}\,\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]^{\gamma}}\Big(K^{\alpha}\,N^{1-\alpha}\Big)^{\gamma}$$

Finally, aggregating over i, yields:

$$Y \;=\; \int y_i\,d\,F(z_i) \;=\; \frac{M\,\mathbb{E}\big[\big((1-\phi_i)^{\gamma}\,z_i\big)^{\frac{1}{1-\gamma}}\big]}{M^{\gamma}\,\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]^{\gamma}}\,(K^{\alpha}\,N^{(1-\alpha)})^{\gamma}$$

so that aggregate productivity is given by:

$$\text{TFP} \;=\; M^{1-\gamma}\Big[\mathbb{E}\big[\big((1-\phi_i)^{\gamma}\,z_i\big)^{\frac{1}{1-\gamma}}\big]\Big]\Big[\mathbb{E}\big[\big((1-\phi_j)\,z_j\big)^{\frac{1}{1-\gamma}}\big]\Big]^{-\gamma}$$

B Additional results

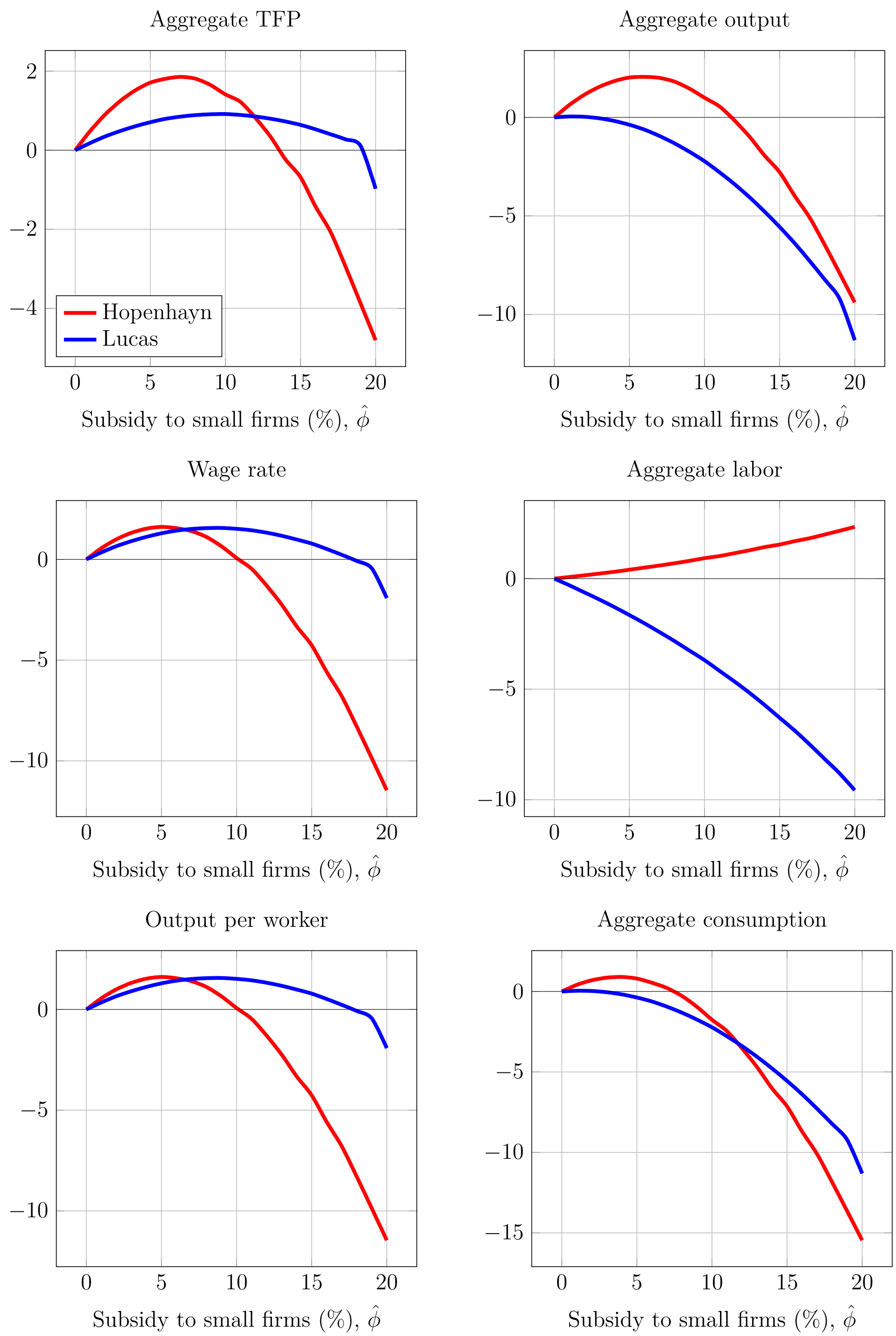

Figure 2: Aggregate impact of distortions (in p.p. relative to baseline)

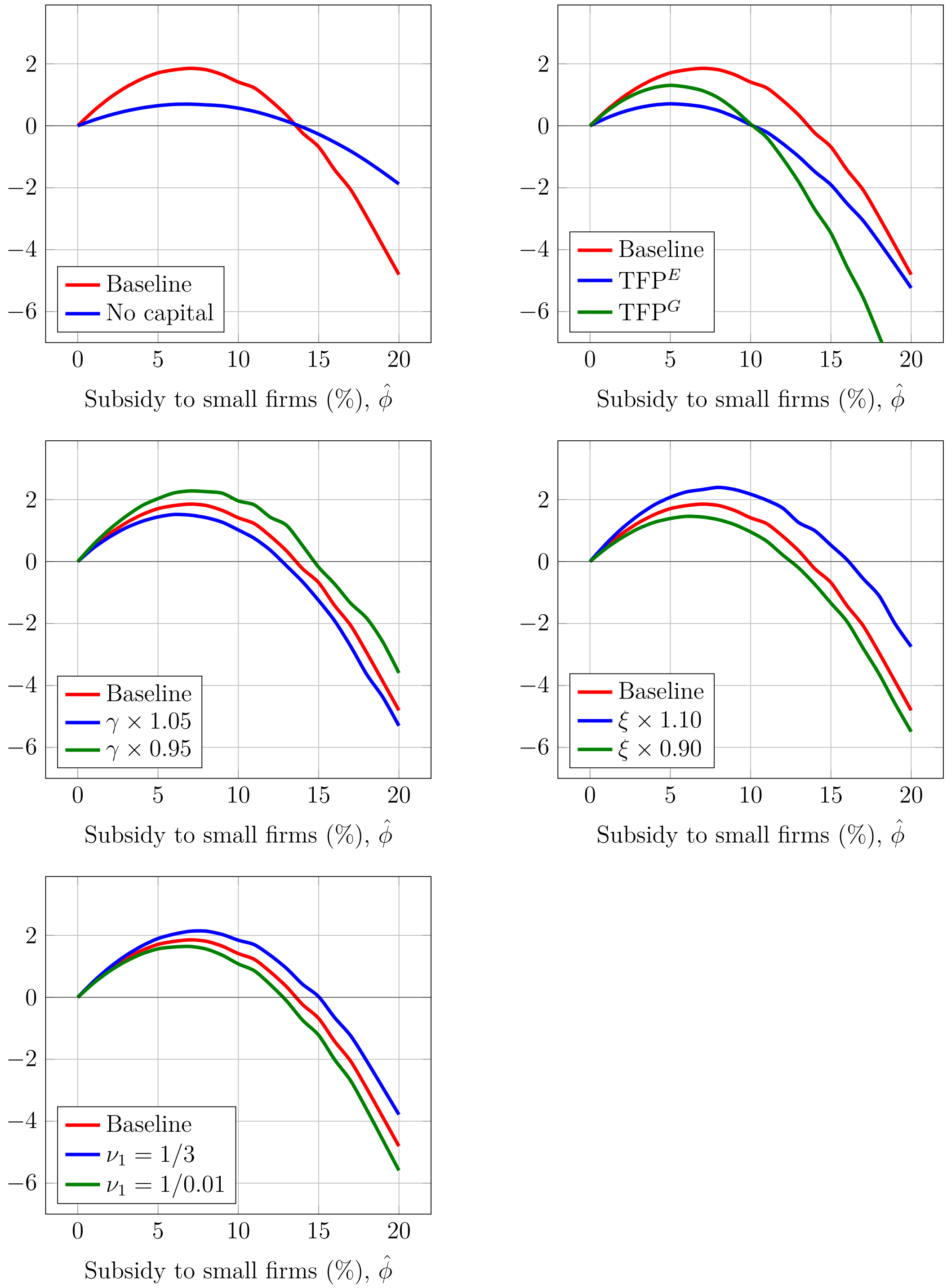

Figure 3: Robustness in the Hopenhayn economy: aggregate TFP impact of distortions (p.p. relative to baseline)

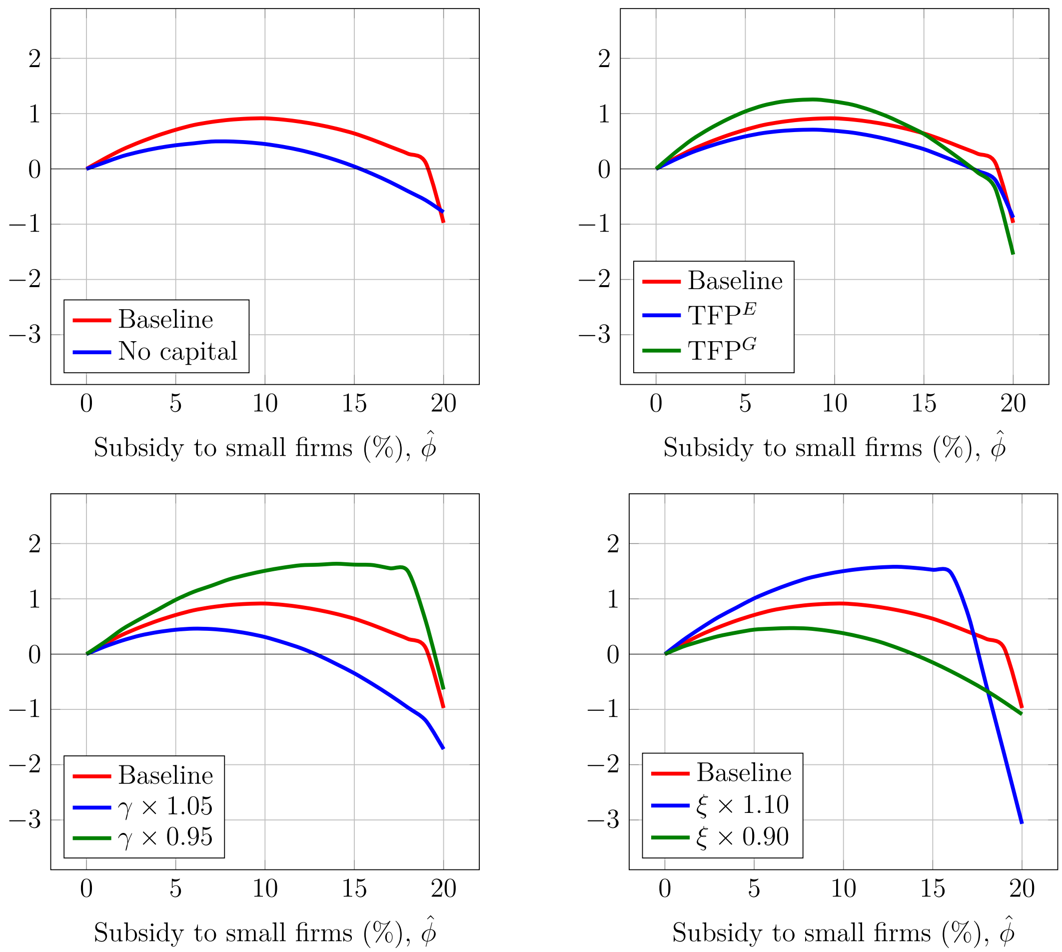

Figure 4: Robustness in the Lucas economy: aggregate TFP impact of distortions (p.p. relative to baseline)