Board of Governors of the Federal Reserve System

International Finance Discussion Papers

ISSN 1073-2500 (Print)
ISSN 2767-4509 (Online)

Number 1393

July 2024

Tax Heterogeneity and Misallocation

Baris Kaymak and Immo Schott

Please cite this paper as:
Kaymak, Baris, and Immo Schott (2024). "Tax Heterogeneity and Misallocation," International Finance Discussion Papers 1393. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/IFDP.2024.1393.

Tax Heterogeneity and Misallocation[*]

Barış Kaymak[†] Immo Schott[‡]

June 25, 2024

Abstract

There is substantial asymmetry in effective corporate income tax rates across firms. While tax asymmetries would reduce productivity in frictionless economies, they can improve efficiency in a distorted economy if taxes alleviate other economic frictions. We develop a framework to estimate to what extent tax asymmetries affect productivity in distorted economies. Using US firm-level balance sheet data alongside measures of effective marginal tax rates, we find a positive correlation between tax rates and factor productivity, suggesting that tax asymmetry exacerbates the distortions from other economic frictions. Eliminating tax rate asymmetries would *raise* aggregate productivity by 3 to 4 percent if taxes distort capital costs alone. Models where taxes also distort the marginal cost of labor predict potential gains as high as 9 percent.

Keywords: Business Taxation, Aggregate Productivity, TFP, Misallocation

JEL Classification: E23, H25, O47

[*]The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

[†]Department of Economic Research, Federal Reserve Bank of Cleveland, 1455 E 6th St, Cleveland, OH 44114, United States. E-mail: barkaymak@gmail.com

[‡]Federal Reserve Board. Division of International Finance, 20th and C Streets, NW, Washington, DC 20551, United States. E-mail: immoschott@gmail.com

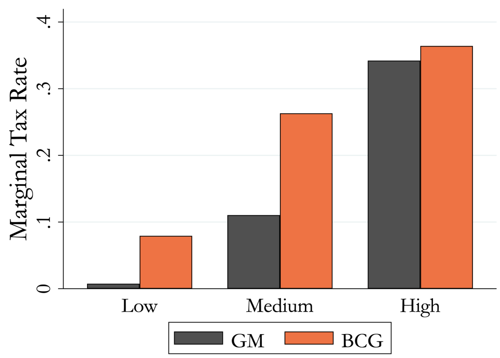
1 Introduction

In an efficient economy, marginal products of inputs are equalized across firms. However, a large number of studies following Hsieh and Klenow (2009) has documented a wide dispersion in marginal productivity across establishments, suggesting that output can be improved by moving resources from firms with low marginal productivity to firms with high marginal productivity. Recent estimates suggest that the US could raise its output by as much as 25 percent if inputs were allocated efficiently (Bils et al., 2021). The sources of input misallocation, however, remain elusive, preventing policy guidance. This paper examines the role of asymmetric corporate taxation as a potential source of input misallocation.

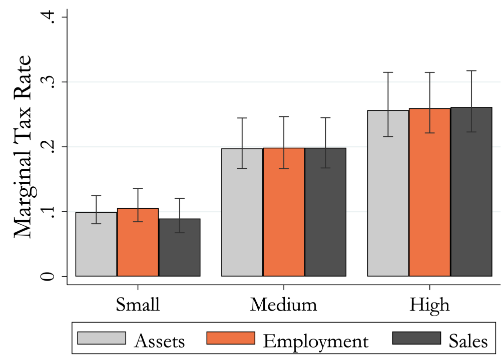
We develop a framework to estimate the effect of tax heterogeneity on aggregate productivity in the US. Although the tax code does not distinguish between individual firms *de jure*, special provisions for deductions and allowances, such as imperfect loss-offsets or the favorable treatment of debt financing, can lead to a dispersion in effective marginal tax rates (EMTR) across firms. Estimates suggest that the resulting dispersion in marginal tax rates is large (Graham and Mills, 2008; Blouin et al., 2010). Figure 1a shows the variation in EMTR in our sample of publicly traded firms in the US between 1980 and 2021. On average, EMTR ranges from under 10 percent for the lowest third of the firms to over 30 percent for the highest third. The standard deviation of tax rates in the cross-section varies between 13 to 17 percent depending on the estimate and year.

Differences in EMTR are systematically related to firm characteristics. Figure 1b shows average EMTR by firm size, measured as total assets, employment, or sales. Larger firms are subject to higher tax rates. Because these firms represent the majority of economic activity, the effect of tax heterogeneity on overall productivity is potentially large, especially if size is an indication of productivity.

Whether these differences in marginal tax rates necessarily worsen the economy's efficiency is not obvious. While heterogeneity in marginal tax rates would lower efficiency in competitive, otherwise frictionless economies, the same cannot be said of economies with distortions (Lipsey and Lancaster, 1956). Efficiency



(a) Tax Rates by Tax Tercile



(b) Tax Rates by Size Class

Figure 1: Asymmetry in Effective Marginal Corporate Tax Rates

Note.- Figure shows average marginal corporate tax rates by terciles of tax rates (panel a), and by terciles of firm size as measured by assets, employment, or sales (panel b). The high and low values shown with whiskers correspond respectively to tax rate estimates from Blouin et al. (2010) and Graham and Mills (2008). Bars show their average. Source: Compustat. Sample period is 1980–2016.

could in fact be *improved* by differential marginal tax rates if they helped offset other distortions in the economy. For instance, the tax advantage of debt lowers the marginal tax of more leveraged firms. This might ease credit frictions and improve investment. Or consider tax carry overs. Losses that are carried over from previous years lower a firm's current marginal tax rate. This might alleviate liquidity constraints and prevent premature firm exit. Whether this is the case is an empirical question.

In pursuit of an answer, we develop a simple theoretical framework that features heterogeneous firms and corporate taxation in a standard model of production. Importantly, while corporate taxes are modeled explicitly, the model includes other, unspecified distortions to factor allocation. We derive analytical formulas which predict the change in total factor productivity resulting from eliminating differences in marginal tax rates across firms. Our results highlight a key statistic which determines the effect of corporate tax policy on efficiency: the correlation between marginal tax rates and factor productivity in the cross-section of firms. This correlation is informative about the interaction between various distortions. A positive correlation implies that firms with higher marginal productivity are taxed more heavily, weakening overall productivity in the economy.

To apply our formulas, we combine balance-sheet data on publicly listed companies in the United States from Compustat with firm-level estimates of the marginal tax rates on corporate income (Graham and Mills, 2008; Blouin et al., 2010). Our baseline calculations indicate that eliminating tax rate heterogeneity would *raise* aggregate TFP in the US by 3 to 4 percent.[1]

Importantly, our results reflect a positive correlation between tax rates and estimated marginal products of capital and labor across firms. This contributes to our estimates in two ways. First, eliminating tax asymmetries lowers tax rates for firms where capital is more productive. This improves TFP by moving capital to those firms. Second, because capital and labor are complementary in production, these firms also raise their employment. The accompanying movement of labor raises productivity even more, because EMTR and labor productivity are also strongly positively correlated in the data. Our calculations suggest that if tax rates were not correlated with other distortions, then potential gains in productivity would be much smaller, less than 1 percent for most years, with the exception of the period prior to the 1986 tax reform. This suggests that corporate tax policy has tended to *exacerbate* distortions to input allocation rather than offsetting them.

The correlation between labor productivity and tax rates is somewhat surprising. Because labor expenses can be deducted from net income, corporate taxes are thought to be non-distortionary to employment. Accordingly, our baseline analysis, where corporate taxes only raise the marginal cost of capital, treats labor distortions as exogenous to taxes. However, taxes can affect the marginal

[1]Following the literature, we focus on static productivity gains from reallocating inputs among existing production units, holding firm-level technologies fixed. Indirect gains from potential improvements in innovation activity, or in the efficiency of firm entry or exit would imply larger gains in the long-run.

cost of labor in models with cash-in-advance requirements for wage payments, employment-based tax credits, or partial expensing of labor costs. In that case, eliminating tax asymmetries mitigates labor distortions, and raises the aggregate TFP over and above our baseline results. Assuming that the correlation between labor distortions and tax rates is *entirely* endogenous, we estimate that tax rate homogenization could improve aggregate TFP by up to 9 percent. We consider this to be an upper bound as the correlation between tax rates and labor productivity is likely not entirely endogenous.

This paper connects to the literature on the macroeconomic effects of factor misallocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). Papers in this literature typically adopt one of two approaches. The first approach consists of measuring the dispersion in marginal products of factors of production and then attributing this dispersion to firm-specific distortions. Comparing the distorted economy to a model-implied frictionless environment, papers following this approach typically find large potential gains from removing frictions (see e.g., Hsieh and Klenow (2009), Gopinath et al. (2017), and Adamopoulos et al. (2022)). A challenge in this literature has been to distinguish between differences in marginal products and differences in production technologies and measurement error (Bils et al., 2021).[2]

The second approach focuses on measuring observable frictions at the firm-level and computing the associated reductions in aggregate productivity. Examples include Gilchrist et al. (2013), Midrigan and Xu (2014), David and Venkateswaran (2019), and Kim (2023). Because of computational and theoretical complexities, this approach is limited in its ability to model and compute a large number of distortions and their interactions. Because specific, observable frictions only explain a fraction of the observed heterogeneity in marginal products across firms,

[2] Hsieh and Klenow (2009) address these issues by benchmarking their misallocation measure to the US when assessing the extent of resource misallocation in India. Bils et al. (2021) use panel data to purge the firm-level dispersion in productivity of measurement error. Because our findings rely on the covariance between tax wedges and productivity, errors in specification or measurement of marginal products are less of a concern. We discuss the implications of measurement error in tax wedges in detail below.

by and large, these studies find the productivity gains from reallocation to be small.

This paper combines elements from both approaches. We focus on a specific friction – corporate income taxes – while explicitly allowing for other, more general, distortions. As in Lipsey and Lancaster's (1956) analysis of the 'second best', removing differences in marginal tax rates across firms in an environment with frictions need not necessarily increase the economy's aggregate productivity, in contrast to most previous work, where eliminating distortions improves productivity by design. The outcome, instead, crucially relies on the joint distribution of tax rates and other distortions. In that sense, our approach is closer to Harberger (1964), and, more recently, Baqaee and Farhi (2020), who study productivity effects of monopolistic price setting in economies with distortions.

Our paper is also related to the literature concerned with measuring the effects of corporate taxes on economic outcomes, such as investment or employment (Harberger, 1962; Hall and Jorgenson, 1967). This literature has focused on the role of specific asymmetries in the tax code, such as those stemming from loss-offset provisions (Auerbach, 1986; Kaymak and Schott, 2019), investment tax credits Cummins et al. (1994), or tax jurisdictions Djankov et al. (2010); Slattery and Zidar (2020) among others. Whereas these papers focus on the effect of the *level* of the corporate tax rate on economic outcomes, we study the *dispersion* in tax rates and its implications for aggregate productivity.

The rest of the paper is organized as follows. Section 2 derives our theoretical results. Section 3 describes how we connect the theory to the data, presents the degree of tax heterogeneity, and shows how tax distortions can be measured. We present our findings in Section 4 and conclude in Section 5.

2 Model

In this section we derive a general formula for quantifying the changes in aggregate TFP from eliminating the heterogeneity in firm-level tax rates in the spirit of Hsieh and Klenow (2009). To fix ideas, we first develop an investment problem with multiple distortions and show how the resulting optimality conditions can be generalized. We also allow for distortions to employment, but maintain an agnostic view of their sources.[3]

Consider the investment problem of a firm that uses capital and labor to produce output with the decreasing-returns-to-scale (DRS) production function, $y = zF(k', n)$, using technology z.[4] The firm faces an effective marginal tax rate on its income, τ, as well as other frictions, such as capital adjustment costs and credit constraints. Each period, the firm chooses capital investment and labor in order to maximize the payout to shareholders, which depends on current and expected after-tax profits. For simplicity, assume that the firm knows the realization of the next period's productivity z and taxes τ at the time of the investment decision. The firm's problem is given by

$$V(z,k) = \max_{n,k',i} \; -i - \Phi\left(\frac{i}{k}\right)k + \beta\left[(1-\tau)\left[zF(k',n) - \omega_n n\right] + E_{z'|z}V(z',k')\right] \tag{1}$$

subject to the law of motion for capital and a collateral constraint

$$k' = i + (1-\delta)k \tag{2}$$

$$i \leq \zeta q k, \tag{3}$$

with associated multipliers q and μ.

The labor decision is not affected by the level of τ because wage payments

[3] In Section 2.2, we consider models where corporate tax rates raise labor costs in addition to the cost of capital and discuss their implications for aggregate productivity.

[4] While we formulate the problem in a perfectly competitive economy with firms facing DRS in production, an equivalent setting would be one where firms have constant returns to scale production, but compete monopolistically as in Hsieh and Klenow (2009).

are tax-deductible.[5] The first-order condition with respect to labor is given by

$$zF_n(k,n) = \omega_n, \tag{4}$$

where ω_n denotes the effective user cost of labor. The optimality condition for the choice of capital is given by

$$q = \beta \left[(1-\tau) zF_k(k',n) - \Phi\left(\frac{i'}{k'}\right) + \Phi'\left(\frac{i'}{k'}\right)\frac{i'}{k'} + q'(1-\delta) + \mu'\zeta' q' \right], \tag{5}$$

with the familiar interpretation that at the optimal choice, the marginal cost of capital, q, equals the discounted future benefit of an additional unit of capital. This benefit consists of the marginal product of capital (net of expected future taxes), the value of non-depreciated capital, and the effects on the marginal costs of investment as well as on the financial constraint.[6]

We define the following two terms that will be used to simplify (5). First, the tax wedge ω_τ is defined as

$$\omega_\tau \equiv \frac{1}{1-\tau}. \tag{6}$$

A higher effective marginal tax rate τ implies a higher value of ω_τ. Second, the "residual" wedge ω_R is defined as

$$\omega_R \equiv \frac{q}{\beta} - q'(1-\delta) - \mu'\zeta' q' + E_{z'|z}\left[\Phi\left(\frac{i'}{k'}\right) + \Phi'\left(\frac{i'}{k'}\right)\frac{i'}{k'} \right]. \tag{7}$$

Any additional frictions a firm might be facing would be included in the residual wedge ω_R. In that sense, the particular distortions firms face other than differential tax rates, be it adjustment costs or credit constraints, do not matter for our results below.[7]

[5] The same is true for capital depreciation and eventual interest payments. Note, however, that we are using *effective* marginal tax rates in our estimations below. These already take into account the tax provisions for depreciation and debt-financing.

[6] The full derivations are relegated to the Appendix.

[7] In what follows, we treat ω_R as exogenous. If ω_R is endogenous to taxes, then eliminating tax rates would also reduce the dispersion in ω_R, and raise aggregate productivity, given the

Using the expression for ω_τ and ω_R, we can write (5) as

$$zF_k(k', n) \quad = \quad \omega_k \quad \equiv \quad \omega_\tau \cdot \omega_R, \tag{8}$$

where ω_k denotes the effective user cost of capital, including all distortions such as capital adjustment costs, financial constraints, and taxes. We decompose ω_k into two parts. The first part, ω_τ, stems from a directly observable distortion (effective marginal tax rates), the second is a residual term, ω_R, that captures all other factors that affect the user-cost of capital.

This formulation allows us to rewrite the firm's problem in (1) as a static allocation problem:

$$\max_{n,k'} -\omega_k k' + zF(k', n) - \omega_n n \tag{9}$$

The first-order conditions of (9) are consistent with the optimality conditions derived in (4) and (8).

2.1 Tax heterogeneity and misallocation

We now derive a formula to measure the effect on aggregate total factor productivity (TFP) stemming from heterogeneity in firm-specific effective costs of capital, ω_K, and labor, ω_N. To do so, we assume a Cobb-Douglas production function $F(k, n) = zk^\alpha n^\beta$. Let $G(\omega_k, \omega_n)$ denote the joint distribution of distortions to capital and labor across firms.

Our aim is to compare total output in the distorted economy with the allocation that a social planner would choose - using the same aggregate quantities of capital $K = \int k \, dG$ and labor $N = \int n \, dG$ as the competitive equilibrium of the distorted economy. Because total inputs are held constant between the two economies, any change in total output $Y = \int y \, dG$ is equivalent to a change in aggregate TFP. We focus on the effect of eliminating tax heterogeneity alone. Therefore we assume that the planner does not (or cannot) change ω_R and ω_n.

---

empirical patterns in the data we document below. In that sense, our findings should be taken to be conservative.

The following proposition gives TFP in the distorted competitive equilibrium of this economy:

Proposition 1. *Total factor productivity in the distorted economy is*

$$Z = \frac{Y}{K^{\alpha}N^{\beta}} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{1-\beta}{1-\gamma}} dG\right]^{\alpha} \left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{1-\alpha}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} dG\right]^{\beta}}. \tag{10}$$

Eliminating tax differentials will generally alter the aggregate demand for capital and labor. We therefore introduce common tax rates (or subsidies) on capital and labor in order to keep the aggregate quantities unchanged. Therefore, the marginal products of labor and capital in the planner's allocations are proportional to ω_R and ω_n, but not equalized across production sites. The optimality conditions in this counterfactual scenario are given by:

$$zF_n(k,n) = \omega_n' = \bar{\omega}_n \cdot \omega_n \quad \text{and} \quad zF_n(k,n) = \omega_k' = \bar{\omega}_k \cdot \omega_R, \tag{11}$$

where $\bar{\omega}_n > 0$ and $\bar{\omega}_k > 0$ represent the planner's tax or subsidy policy that is common across firms. The resulting optimality conditions for k and n coincide with those of a profit-maximizing firm that takes the distortions and the common tax wedges $\bar{\omega}_k$ and $\bar{\omega}_n$ as given. Those wedges are chosen to satisfy the input-neutrality constraints on the allocation problem:

$$\int k(\omega_n', \omega_k') dG' = K \text{ and } \int n(\omega_n', \omega_k') dG' = N$$

where $G'(\omega_k', \omega_n')$ denotes the distribution associated with the new distortions. Because wedges that are common to all firms do not distort relative marginal products, they do not cause a misallocation of inputs in the cross-section of firms.[8] Consequently, $G'(\omega_k', \omega_n')$ is equivalent to $G'(\omega_R, \omega_n)$ in terms of its im-

[8]Note that because this does not eliminate corporate taxation altogether, aggregate capital investment remains suboptimal overall.

plications for TFP distortions.

The counterfactual TFP in the absence of tax heterogeneity can be obtained by setting $\omega_\tau = 1$ (or to any positive scalar), and replacing ω_k by ω_R in equation (10) as ω_R is the only remaining determinant of the marginal cost of capital. The resulting counterfactual TFP is given in the proposition below.

Proposition 2. *Total factor productivity in the counterfactual economy with homogeneous tax rates across firms is*

$$Z^* = \frac{\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{1-\beta}{1-\gamma}} \omega_\tau^{\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{1-\alpha}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\alpha}{1-\gamma}} dG\right]^\beta}. \tag{12}$$

The ratio of the counterfactual TFP in equation (12) to the actual TFP in equation (10) gives the marginal effect of eliminating tax heterogeneity on aggregate productivity. From equation (12), $Z^* = Z$ when ω_τ is constant across firms and $Z^* > Z$ whenever ω_τ is heterogeneous but orthogonal to other distortions.[9] More generally, however, Z^* can be higher or lower than Z, depending on how tax rates correlate with other distortions across firms.[10]

To gain further insights, let us now consider an economy where the distortions are distributed jointly according to a log-normal density. The resulting formulas help form an intuition about the sources of misallocation from tax heterogeneity. They show how interactions between tax rates and other distortions can play an important role in assessing the allocative effects of tax distortions. We relax the log-normality restriction in our quantitative analysis below. Under that assumption, TFP in the distorted economy is equivalent to:

$$\ln Z = \ln \int z^{\frac{1}{1-\gamma}} - \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)\sigma_k^2 + \beta(1-\alpha)\sigma_n^2 + 2\alpha\beta\sigma_{kn}\right], \tag{13}$$

where σ_k^2 and σ_n^2 denote the variances of $\ln \omega_k$ and $\ln \omega_n$ respectively, and σ_{kn}

[9]This follows from Jensen's inequality combined with the fact that $1-\beta > \alpha$.

[10]For instance, $Z^* < Z$ when $\omega_R = 1/\omega_\tau$ and $\omega_\tau \perp \omega_n$.

is the covariance between them. Aggregate TFP reflects the underlying distribution of micro-level productivity levels, z, adjusted for efficiency losses caused by input distortions. We are interested in the change in TFP from eliminating the heterogeneity in tax rates. This implies setting ω_k' to $\bar{\omega}_k \cdot \omega_R$, i.e., eliminating the tax wedge from the marginal product of capital in (8). The counterfactual TFP, $\ln Z^*$, can be obtained by setting $\sigma_k^2 = \sigma_R^2$ and $\sigma_{kn} = \sigma_{Rn}$ in equation (14). Taking differences and rearranging terms gives the next proposition.

Proposition 3. *If ω_k, ω_n, and ω_τ are jointly log-normally distributed, then eliminating the heterogeneity in the marginal tax rates yields the following change in aggregate TFP:*

$$\ln \frac{Z^*}{Z} = \frac{1}{2}\frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2 + \frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1) + \frac{\alpha\beta}{1-\gamma}\sigma_\tau^2 L_{n\tau}, \tag{14}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_k$ and $\ln \omega_n$ on $\ln \omega_\tau$.

Equation (14) presents a simple way of capturing the total change in TFP from eliminating tax heterogeneity. It has three distinct components. The first one is a pure misallocation component, representing the reduction in aggregate output caused by a dispersion in the marginal products of capital across firms. In the absence of other economic distortions, or if such distortions were orthogonal to tax rates, this would be the total improvement in TFP that can be expected from equalizing marginal tax rates. The magnitude of the TFP gains is increasing in the variance of the tax rates, σ_τ^2, the span of control parameter, γ, and capital's share of income, α.[11]

The second term in (14) captures the correlation of marginal tax rates with other distortions to capital. The coefficient $L_{k\tau}$ is less than one if $\ln \omega_R$ and $\ln \omega_\tau$ are negatively correlated. This can arise if taxes alleviate other capital distortions. Eliminating tax differentials in this case need not lead to TFP gains. In fact, in the

[11]To see this, note that $\frac{\alpha(1-\beta)}{1-\gamma} = \alpha + \frac{\alpha^2}{1-\gamma}$.

extreme case where tax rates fully offset other distortions, $\omega_\tau = \omega_R^{-1}$, $L_{k\tau} = 0$, which implies that the first two terms cancel each other.

Of course, taxes could also exacerbate the existing distortions. If firms that have higher marginal costs of borrowing also have higher marginal tax rates, this would manifest as $L_{k\tau} > 1$, and raise the potential gains from eliminating tax differentials.

The last term in (14) captures the correlation of marginal tax rates with distortions to labor. Equalization of tax rates lowers the tax rates for some firms and raises their input demand resulting in a reallocation of labor toward those firms. This improves efficiency only if the marginal product of labor, ω_n, is higher at those firms on average, i.e., if $L_{n\tau} > 0$.

2.2 Alternative models of labor distortions

In our equations above, we treated ω_n as exogenous to ω_τ, because basic theory suggests that corporate tax rate should not distort employment, conditional on capital and firm-level TFP. This follows from the fact that outlays on labor are typically deducted from the tax base. There are nonetheless situations where corporate tax rates may have a direct effect on employment decisions, for instance, when part of the labor cost cannot be deducted from income or when there are explicit employment-based credits and incentives, such as the recent employment retention credit provided by the CARES act. Similarly, models with cash-in-advance requirements, where wages have to be paid before sales are realized, also lead to a tax-related wedge in the optimality condition for employment. In this subsection, we consider models with imperfect expensing of labor costs. In the Appendix, we present a cash-in-advance model and illustrate how corporate tax rates can directly affect employment decisions in addition to their indirect effect through investment.

If there is an endogenous relationship between corporate taxation and the marginal cost of labor, then eliminating tax differentials would also affect labor distortions. This would potentially lead to larger TFP gains than implied by

Proposition (3) if taxes raise the marginal cost of labor. Ascertaining the causal effect of taxes on labor costs is not straightforward as it requires estimates of how taxes alter firm-specific labor supply schedules. Some progress can be made, however, on the *potential* role of these alternative models by interpreting the empirical correlation between labor productivity and tax rates as entirely endogenous. As we demonstrate next, this provides an upper bound on the productivity gains from homogenizing tax rates across firms when taxes raise marginal labor costs.

Consider a scenario where a fraction $\lambda \in [0,1]$ of labor costs are not deductible from the corporate tax base. Denoting the total cost per employee by ω_l, the after-tax income is now given by $(1-\tau)[zF(k',n)-(1-\lambda)\omega_l n]-\lambda\omega_l$, which leads to the following optimality condition for employment:

$$zF_n(k,n) = (1-\lambda+\lambda\omega_\tau)\omega_l = \omega_n \tag{15}$$

When $\lambda > 0$, a higher corporate tax rate raises the marginal cost of labor. Importantly, heterogeneity in tax rates can lead to a dispersion in marginal costs of labor ω_n across firms, even when all firms face the same total cost per employee ω_l.[12] Unlike our baseline model above, eliminating tax heterogeneity in this economy reduces the dispersion in marginal cost of labor, thereby leading to larger gains in aggregate productivity.

In our associated calculations below, we consider the extreme case of $\lambda = 1$, which implies $\omega_n = \omega_l \cdot \omega_\tau$. This is tantamount to interpreting the correlation between labor productivity and tax rates as structural. The following proposition gives the counterfactual TFP from eliminating tax heterogeneity in that scenario:

Proposition 4. *In models with imperfect labor expensing, total factor productivity*

[12] A similar distortion can also arise when $\lambda < 0$, i.e., when there is a tax credit for employment. We focus on $\lambda > 0$ here given the positive empirical correlation between labor productivity and tax rates below.

in the counterfactual economy with homogeneous tax rates across firms is:

$$Z^{**} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\gamma}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{1-\beta}{1-\gamma}} \omega_\tau^{\frac{1}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{1-\alpha}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{1}{1-\gamma}} dG\right]^\beta}. \tag{16}$$

Dividing Z^{**} above with Z in equation (10) gives the relative aggregate changes in TFP from homogenizing tax rates across firms. When all distortions are distributed log-normally, the implied change in aggregate TFP can be simplified as follows.

Proposition 5. *Suppose $\omega_k = \omega_R \cdot \omega_\tau$, $\omega_n = \omega_l \cdot \omega_\tau$, and ω_τ are jointly log-normally distributed. Eliminating the heterogeneity in marginal tax rates ($\sigma_\tau = 0$) yields the following change in aggregate TFP in models with imperfect labor expensing:*

$$\ln \frac{Z^{**}}{Z} = \frac{1}{2} \frac{\gamma}{1-\gamma} \sigma_\tau^2 + \frac{\alpha}{1-\gamma} \sigma_\tau^2 (L_{k\tau} - 1) + \frac{\beta}{1-\gamma} \sigma_\tau^2 (L_{n\tau} - 1), \tag{17}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_k$ and $\ln \omega_n$ on $\ln \omega_\tau$.

Comparing equations (14) and (17) reveals the significance of a structural link between taxes and labor costs. The first component, attributable to the dispersion in the tax rates is larger in (17), because it now includes the direct contribution of tax rate heterogeneity to the dispersion in marginal cost of labor, ω_n. The second component captures the correlation between tax rates and other distortions to capital. Because taxes now raise marginal costs of both capital and labor, they generate a positive correlation between the two. This exacerbates the distortionary effect of taxes (via σ_{kn} in equation (13)). The last component in (17) can be smaller or larger than in equation (14) because it now captures the productivity implications of the correlation between the tax wedge, ω_τ, and *other* labor distortions, ω_l.

In what follows, we present two sets of results: the baseline scenario, where we interpret the empirical relation between labor distortions and tax rates as non-structural as in Section 2.1, and an alternative scenario, where we interpret it as structural. We consider the latter as an upper bound on the allocative effects of tax heterogeneity.

Next we describe the data and our measurement methodology.

3 Data and methodology

In this section we discuss the data sources and potential issues with the measurement of effective marginal corporate income tax rates. We then describe our methodology and examine the patterns of correlations between tax rates and firm-level productivity, which serve as inputs to our formulas above.

3.1 Data sources and definitions

Our main data source is the Compustat database covering the years from 1980 to 2021. Compustat provides annual balance-sheet data on publicly listed companies. To conduct our calculations we use information on output, employment, and the capital stock.[13] We define output as the sum of sales and changes in inventories during the year.[14] We measure labor input by employment. To construct a measure of a firm's capital stock, we use a perpetual inventory method using investment expenditures. This allows us to compute the average productivity of labor and capital for each firm and year.

We supplement this data with estimates of firms' marginal corporate income tax rates, taken from two sources: Graham and Mills (2008) and Blouin et al. (2010). These studies take into account such factors as loss-offset provisions, depreciation allowances, and debt service when calculating an effective rate for each firm. Graham and Mills (2008) and Graham (1996) show that the simulated

[13] See Appendix B for details on the data used.

[14] Compustat does not contain information on the cost of intermediate inputs, preventing a measure of value added.

tax rates provide a close approximation of the actual taxes paid as reported in tax records.

Table 1: Summary statistics of marginal tax rates

Variable	mean	sd.	p25	p50	p75	N
τ^{GM}	0.169	0.171	0.007	0.070	0.342	125.048
τ^{BCG}	0.240	0.136	0.108	0.283	0.342	159.247

Note.– The two effective tax variables are taken from Graham and Mills (2008) and Blouin et al. (2010). See Appendix D for variable definitions and sample selection.

Summary statistics for the two effective marginal tax rate measures are shown in Table 1. The two tax measures differ somewhat in methodology. Rates estimated by Graham and Mills (2008) show more bunching at zero and at the top statutory marginal tax rate, which has varied over the years. Rates estimated by Blouin et al. (2010) provide a smoother distribution, with a higher average rate and a slightly smaller variance. The two tax measures are highly but imperfectly correlated ($\rho = 0.61$).

3.2 Measuring tax distortions

Because we have two distinct tax rate measures that are correlated imperfectly, we treat each measure as an erroneous estimate of the true marginal tax rate. This allows us to leverage the empirical content of each measure by focusing on their common component. Specifically, we interpret each measure as a combination of the true marginal tax rate and a classical measurement error:

$$\ln \omega_{i\tau}^{*} = \ln \omega_{\tau} + \epsilon_i,$$

where ϵ_i is measurement error with variance $\sigma_{\epsilon,i}$ and conditional mean $\mathbb{E}[\epsilon_i|\omega_\tau] = 0$ for $i \in \{1, 2\}$. Replacing the tax wedges by their measured counterparts not only biases the estimates of the total allocative effect of tax heterogeneity, but it can also lead to a misinterpretation of how tax rates interact with capital and labor distortions.

To gain intuition, let us first consider equation (14), which we use for our decomposition exercises below, before we turn to the implications of measurement error for the general, non-linear formula. Equation (14) has three empirical moments that depend on the tax measure: the variance of the tax wedge, σ_τ^2, as well as the interactions of tax wedges with capital and labor productivity, summarized by the projection coefficients $L_{k\tau}$ and $L_{n\tau}$. When ω_τ is replaced by its measured counterpart, ω_τ^*, all three moments are estimated with a bias. The variance of ω_τ^* is inflated relative to the variance of ω_τ by a factor of $(\sigma_\tau^2 + \sigma_\epsilon^2)/\sigma_\tau^2$. A larger measured dispersion in tax wedges tends to exaggerate the magnitude of the estimated change in TFP. On the other hand, the estimates of $L_{k\tau}$ and $L_{n\tau}$, i.e., the projections of capital and labor productivity on tax wedges, are attenuated proportionally by $\sigma_\tau^2/(\sigma_\tau^2 + \sigma_\epsilon^2)$ when ω_τ^* is used. Lower measured correlations between tax wedges and other distortions tend to attenuate the measured change in TFP from eliminating tax heterogeneity. The net effect of these two forces is a downward bias in the estimated gains as summarized by the following proposition.

Proposition 6. *Assume that the tax wedge is measured with error, $\omega_\tau^* = \omega_\tau + \epsilon$ with $\mathbb{E}(\epsilon|\omega_\tau) = 0$. Then, replacing the ω_τ by ω_τ^* in equation (14) underestimates the net TFP gain from eliminating tax heterogeneity:*

$$\ln(Z^*/Z)|\omega_\tau^* = \ln(Z^*/Z)|\omega_\tau - \frac{\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}. \tag{18}$$

Additionally, the attenuated estimates of $L_{k\tau}$ and $L_{n\tau}$ result in an inaccurate assessment of the interaction between tax wedges and other capital distortions. The capital component (second term in equation (14)) gets attenuated, and the direct component (first term) gets exaggerated. The labor component (third term) is unaffected because the attenuation bias when estimating $L_{n\tau}$ is offset by the upward bias when estimating σ^2_τ.[15] [16]

We address measurement error when we use equation (14) as follows. First, we estimate the variance of tax wedges with the covariance of the two measures: $\hat{\sigma}^2_\tau = cov(\ln \omega^*_{1\tau}, \ln \omega^*_{2\tau})$. Second, we estimate $L_{n\tau}$ and $L_{k\tau}$ by regressing labor and capital productivity on the measured tax wedges using an instrumental variables approach, where each tax measure is used as an instrument for the other. This method is known to yield consistent estimates when the errors are uncorrelated across measures. Specifically, for each measure $i, j \in \{1, 2\}$ with $i \neq j$, $\hat{L}^{IV}_{i,k\tau} = cov(\ln \omega^*_{i\tau}, \ln k)/cov(\ln \omega^*_{i\tau}, \ln \omega^*_{j\tau})$. $\hat{L}^{IV}_{i,n\tau}$ is defined similarly. We substitute these estimates in equation (14) to compute the change in aggregate TFP and its components.

However, because the distribution of tax rates shown in Table 1 is bimodal, it departs from a log-normal distribution. Therefore, we also use the generalized formulas in equations (10), (12) and (16) to compute the potential productivity gains. In particular, we partition the firms into equally sized quantile bins based on their (measured) marginal tax rate in each year. For each group-year cell, we then calculate the average tax wedge and compute the average firm-level TFP, along with capital and labor productivity as described above. In the Appendix, we show that the resulting bias for measurement error aligns with the error in equation (18) when the distribution of marginal products are log-normal, conditional on the tax rate, regardless of the marginal distribution of the tax rates. Accordingly, we adjust the resulting estimates of TFP gains for measurement error using equation (18). This requires an estimate for the variance of the measurement error, which we estimate by subtracting the covariance between the two measures, our estimate for the true variance, from the total variance of the measured wedge $\hat{\sigma}^2_{\epsilon,jt} = var(\ln \omega^*_{j\tau}) - cov(\ln \omega^*_{1\tau}, \ln \omega^*_{2\tau})$ for $j \in \{1, 2\}$. Because we are interested in the cross-sectional dispersion in tax wedges, we repeat this

[15] This follows from the usual attenuation bias formula: $E[\hat{L}^{OLS}_{n\tau}] \times \sigma^2_{\omega^*_\tau} = L_{n\tau} \times \sigma^2_{\omega_\tau}$, where $\hat{L}^{OLS}_{n\tau}$ is the OLS coefficient obtained by regressing $\ln \omega_n$ on $\ln \omega^*_\tau$

[16] The corresponding bias from replacing ω_τ with ω^*_τ in equation (17), i.e., When the correlation between labor distortions and tax rates is causal, is $-\frac{\gamma}{1-\gamma}\frac{\sigma^2_\epsilon}{2}$.

for each year.

3.3 Correlations between productivity and tax distortions

To measure the distortions to capital and labor, we use the optimality conditions for factor demands, $\ln \omega_k = \ln \alpha + \ln(y/k)$ and $\ln \omega_n = \ln \beta + \ln(y/n)$, where factor shares α and β are common to all firms in an industry during a given year. To estimate the correlation between total distortions to capital and the tax wedge, we estimate the specification

$$\ln(y/k)_{it} = D^k_{st} + L_{k\tau,t} \ln \omega^*_{\tau,it} + e^k_{it}, \tag{19}$$

where i denotes the firm, and t the year of observation. D^k_{st} are indicators for a full set of sector and year interactions. These indicators capture variations in capital shares and average distortions across sectors and years. Therefore, $L_{k\tau,t}$ reflects the correlation between the tax wedge and other capital distortions across firms in a given year, that is, our estimate of $L_{k\tau}$ in equation (14).

We estimate the cross-sectional correlation between distortions to labor and tax wedges using a similar specification:

$$\ln(y/n)_{it} = D^n_{st} + L_{n\tau,t} \ln \omega^*_{\tau,it} + e^n_{it}, \tag{20}$$

Because each tax estimate might contain measurement error, the OLS estimates of $L_{k\tau}$ and $L_{n\tau}$ are potentially attenuated. To remedy this issue, we estimate equations (19) and (20) with an instrumental variables approach, where one tax measure is used as an instrument for the other.

We estimate separate values of $L_{k\tau,t}$ and $L_{n\tau,t}$ for each year to compute TFP gains or losses below. In Table 2 we summarize the patterns of correlations between different distortions using a common estimate for all years. The first three columns show the OLS estimates of (19) and (20). The first row in each column shows the estimates obtained by tax measures provided by Graham and

Mills (2008), and the second row shows the corresponding estimate using the tax measures from Blouin et al. (2010).

Both measures suggest a strong positive correlation between taxes and capital productivity (column 1), labor productivity (column 2), or firm-level TFP (column 3), although they somewhat disagree on the strength of that correlation.

The two measures in the first column imply different correlation patterns between tax wedges and other distortions to capital. Recall that a value below (above) one indicates that the tax wedge is negatively (positively) correlated with other distortions to capital: $cov(\ln \omega_R, \ln \omega_\tau) < 0$ (> 0). Therefore while Graham and Mills's estimates of the EMTR suggest a negative correlation between the tax wedge and capital distortions, Blouin et al.'s estimates suggest they are orthogonal. By contrast, both measures imply a positive correlation between distortions to labor and the tax wedge in column 2: $cov(\ln \omega_n, \ln \omega_\tau) > 0)$.

Table 2: Tax distortions and factor productivity

	(1) $\ln y/k$	(2) $\ln y/n$	(3) $\ln z$	(4) $\ln y/k$	(5) $\ln y/n$	(6) $\ln z$
			Panel A			
$\ln \omega_{1\tau}$	0.67	0.55	0.65	1.26	1.61	1.72
	(0.04)	(0.03)	(0.02)	(0.08)	(0.06)	(0.05)
			Panel B			
$\ln \omega_{2\tau}$	0.96	1.23	1.30	1.59	1.31	1.56
	(0.06)	(0.04)	(0.04)	(0.09)	(0.06)	(0.06)

Note.– The table shows the results from regressions of productivity on the tax wedge $(1/(1-\tau))$. Columns 1 to 3 report OLS estimates. Columns 4 to 6 report instrumental variable estimates to correct for measurement error. All specifications control for a full interaction of sector and year indicators. Data on productivity come from authors' calculations from Compustat. Data on marginal tax rates come from Graham and Mills (2008) in Panel A and Blouin et al. (2010) in Panel B.

Correcting for the attenuation bias yields a more consistent picture across measures as shown in columns 4 to 6. Both estimates in column 4 indicate that other distortions to capital correlate positively with the tax wedges across firms, implying that the heterogeneity in tax rates exacerbates the existing distortions to capital.

The estimates in column 5 imply a positive correlation between labor distortions and the tax wedge. The two alternative models presented in Section 2 interpret that correlation differently. From the perspective of our baseline model in Section 2.1, this is surprising. Because labor costs are deducted from the tax base, employment decisions should not in principle be affected by corporate taxes, and the relationship between the two wedges is therefore considered spurious. In the model presented in Section 2.2, on the other hand, higher tax rates directly raise the marginal cost of labor. Therefore, a positive correlation is expected.

Columns 3 and 6 show the projections of firm-level TFP on the tax wedge, which we use below for some of our results. Consistent with the patterns from average labor and capital productivity, these estimates show that productive firms on average face higher marginal tax rates.

In our calculations below, we compute the IV estimates for each year. Because from the two tax measures we obtain two estimates of the same underlying parameter, we combine them by taking an average, weighted by the inverse of their variance. For example, using the estimates for the entire sample period, reported in columns 4 and 5 of Table 2, we obtain a value of 1.41 (s.e. 0.04) for $L_{k\tau}$ and 1.38 (s.e. 0.02) for $L_{n\tau}$ for the entire sample period.[17]

Before we turn to the implications of our estimates for aggregate TFP and output, we need to set values for the parameters of the production function. For our baseline results, we keep those parameters fixed over time. This allows us to

[17] Specifically, the variance minimizing weights are $\lambda_1 = var(b^{IV}_{k\tau,2})/(var(b^{IV}_{k\tau,1})+var(b^{IV}_{k\tau,2}))$ for $b^{IV}_{k\tau,1}$, and $\lambda_2 = 1 - \lambda_1$ for $b^{IV}_{k\tau,2}$. For $L_{k\tau}$, for instance, the values in column (4) of Table 2 imply a weight of $0.44 = 0.08^2/(0.08^2 + 0.09^2)$ for 1.59 and 0.56 for 1.26, giving a weighted average of 1.41. The standard error for the weighted estimate is given by $var(b^{IV}_{k\tau,1}) \times var(b^{IV}_{k\tau,2})/(var(b^{IV}_{k\tau,1}) + var(b^{IV}_{k\tau,2}))$.

highlight the changes in the interactions between distortions when presenting the time trends. To that end, we set $\gamma = 0.85$, which implies a profit share of 15 percent in output. We set $\beta = 0.85 \times 2/3$, which gives a labor share of 0.57. These two choices imply a capital share of $\alpha = 0.28$. Given the downward trend in the labor share of income during our sample period, we also consider alternative scenarios for these parameters and discuss their implications for our findings below.

4 Results

In this section, we report our estimates for the effect of eliminating cross-sectional differences in tax wedges on total factor productivity. Our primary finding is that eliminating tax heterogeneity would improve TFP in the United States. Our baseline estimates put this gain between 3 and 4 percent. We find that this result is robust to including macroeconomic trends in the decline of the labor share or an increase in markups. We then decompose the potential TFP gains associated with eliminating tax heterogeneity into its three components. The pure dispersion in tax rates explains a sizeable part of the estimated TFP gains, but this component became smaller after the tax reforms of the 1980s. The correlation of tax rates with capital and especially labor make up the largest component of projected gains from tax misallocation.

4.1 Aggregate productivity

We begin with the estimates of the overall TFP gains associated with eliminating tax heterogeneity shown in Figure 2. The solid black line shows our baseline estimate. The change in TFP from eliminating tax heterogeneity is positive throughout our sample period. This suggests that differential tax rates do not offset other distortions to capital and labor. On the contrary, our estimates imply that factors are likely more productive at firms that have higher marginal tax rates. The average gain in TFP across all years is 3.3 percent. Given the TFP growth in the

US during our sample period, this corresponds to four years of average annual TFP growth. Over the sample period, the potential TFP gain varies between 2 percent and 5 percent, but there is no discernible time trend.

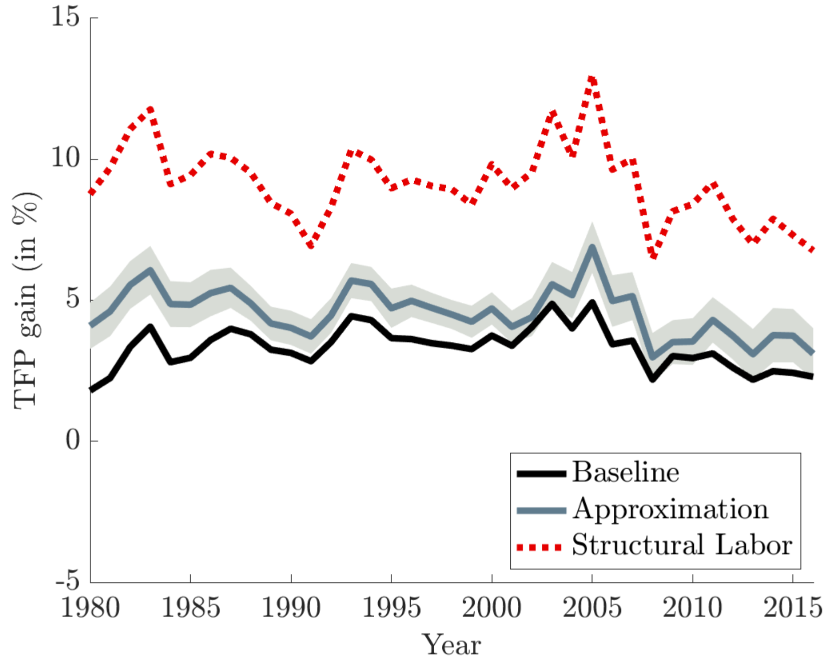


Figure 2: TFP gains from tax rate equalization

Note.– The figure shows the estimated change in aggregate TFP if all firms were to face a common corporate tax rate. Baseline model (and its approximation) assume corporate taxes only distort investment. The structural labor model additionally allows taxes to distort hiring decisions.

To understand the sources of the productivity gain, we now turn to our approximation of the TFP gains from tax homogenization in equation (14). First, we assess the accuracy of the approximation by comparing the implied total productivity gain with our finding above.

The gray line in Figure 2 shows the estimated total TFP gains using the approximation formula. The shaded gray area represents error bands corresponding to two standard errors above and below the point estimate (associated with the standard errors of $L_{k\tau}^{IV}$ and $L_{n\tau}^{IV}$). Using the approximation, the magnitude of the gain is 4.6 percent on average. This reinforces our main finding that that tax heterogeneity reduces aggregate TFP. Although slightly larger than the baseline estimate, the estimates lie reasonably close to our baseline results. The two series have a very high correlation coefficient of ($\rho = 0.81$). The approximate gains range from around 5 percent during the early 1980s to around 3 percent in recent years, suggesting a 2 percent improvement over our sample period.

Next, we study the different components of the approximated gain in order to understand the interaction of tax heterogeneity with other distortions in the economy.

From equation (14) the total TFP gains reflect not only the dispersion in tax rates, but also the estimated correlation between the tax wedge and other distortions to capital and labor. The stacked bars in Figure 3 decompose the total TFP gains into these three components. The solid line in Figure 3 gives the total TFP gain and is identical to the gray line in Figure 2.

The blue bars reflect the misallocation caused purely by tax heterogeneity. If tax wedges were uncorrelated with the marginal products of capital and labor, the blue bars would represent the total gains in TFP. They can therefore be interpreted as the TFP distortion caused by tax heterogeneity in an otherwise frictionless economy. From Figure 3 this component of the TFP gain was 1 to 2 percent prior to 1986, due mainly to the higher statutory corporate tax rate, and has been under 1 percent since then. Overall, it represents less than a quarter of the total TFP gains. This highlights the importance of taking other distortions into account when studying the allocative effects of a particular distortion.

The red bars in Figure 3 represent the TFP gains that stem from how tax rates correlate with capital productivity. There are two hypothetical cases. If $cov(\omega_R, \omega_\tau) < 0$, or equivalently, $L_{k\tau} < 1$, the heterogeneity in tax rate *reduces* the distortionary effects of other distortions. This could be the case if firms that face relatively large distortions, for example, due to credit constraints or adjustment costs, face *lower* tax rates. In that situation, which is observed for several

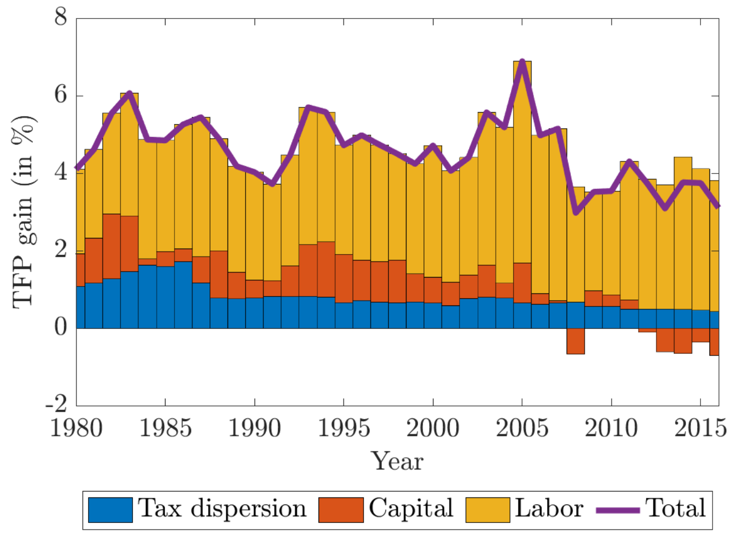


Figure 3: Components of TFP gains from eliminating tax heterogeneity

Note.– The figure shows the three components of TFP gains from tax rate equalization across firms (see equation (14)). The blue bars labeled "Tax dispersion" represent gains in an otherwise frictionless economy. The red bars labeled "Capital" show additional gains/losses in an economy with capital distortions. The bars labeled "Labor" show additional gains in an economy with capital and labor distortions.

years in the figure, the red bars contribute negatively to the TFP gain. The second case, i.e., when $cov(\omega_R, \omega_\tau) > 0$ is the more common case, however. This implies that tax rates are positively correlated with capital distortions. From the figure this correlation is not very strong, however, resulting in less than 1 percent of additional TFP gains. Over the years, the potential gains implied by this component have been highest during the 1990s and early 2000s and have diminished in more recent years. This suggests an improvement from an efficiency perspective in the distribution of tax rates across firms over time.

The third component of the TFP gain comes from the correlation between labor distortions and tax wedges. This is shown as the yellow bars in Figure 3.

When that correlation is positive, equalizing taxes results in lower tax rates for firms where the marginal product of labor is typically higher. Because capital and labor are complementary in production, a lower tax rate also results in reallocation of employment toward high marginal product firms and creates an additional gain in TFP. Quantitatively, those gains represent a majority of the total gains depicted in Figure 3. This is partly because the share of labor in total income is roughly twice as large as that of capital.[18]

Overall, tax heterogeneity alone represents less than 1 percent of the total gains on average, across the years. The majority of the projected gain in TFP from tax homogenization is due to the fact that tax wedges are correlated positively with other labor distortions.

4.2 Capital accumulation and potential output gains

The estimated TFP gains above are defined as gains in total output for given quantities of aggregate labor and capital. In the long-run, however, the gain in productivity can cause additional capital accumulation and lead to additional output gains. To gauge the magnitude of this effect, consider the standard, representative agent growth model where labor is supplied inelastically. There, a 1 percent increase in TFP implies a $1/(1-\alpha)$ percent increase in the capital stock: $K = (\alpha Z/r)^{\frac{1}{1-\alpha}}$, where Z is aggregate TFP. This raises total output by $\alpha/(1-\alpha)$ percent, because α is the capital share in production. This increase occurs in addition to the 1 percent increase in output due to higher TFP. In our baseline calibration, we assumed profits represent 15 percent of output and capital's cost share is $1/3$, implying $\alpha = 0.85/3$. These assumptions suggest an additional gain of about 0.4 percent in total output from capital accumulation in the long run. For our baseline estimate of 3.3 percent TFP gain from tax homogenization, this

[18] More generally, capital and labor co-move during the reallocation from eliminating tax heterogeneity as long as the elasticity of substitution between capital and labor is not too high (less than $1/(1-\gamma)$).

translates to a 4.6 ($= 3.3 \times 1.4$) percent long-run gain in total output per capita.[19]

A related consequence of output growth is higher corporate tax revenue, which allows for a lower overall tax rate and further increases output growth through capital accumulation. Revenue-neutral gains in output are therefore larger than input-neutral gains. Specifically, revenue neutrality requires that the tax rate is reduced to offset the percent gain in output because net corporate income is proportional to output in our model: $\Pi = (1-\gamma)Y$. Each percent decline in the tax rate raises the capital stock by $\frac{1}{1-\alpha}\frac{\tau}{1-\tau}$, and output by α times the capital stock.[20] Additional output allows for even lower tax rates, and so on. The associated multiplier is $\left(1 - \frac{\alpha}{1-\alpha}\frac{\tau}{1-\tau}\right)^{-1}$, which is about 1.1 when $\alpha = 0.85/3$ and the (common) tax rate is 20 percent, roughly the average effective tax rate in our sample between the two estimates. Therefore, a revenue-neutral tax homogenization reform would raise output per capita by 5.1 percent (4.6×1.1) in the long-run, once capital accumulation is accounted for.

Selection of firms through entry and exit is another important margin whereby tax homogenization can improve aggregate productivity. In standard models of industry dynamics without distortions (e.g. Hopenhayn (1992)), efficiency requires the survival of firms with higher productivity. Tax heterogeneity can distort that selection by allowing less productive firms to survive longer and potentially pushing productive firms into premature exit because tax rates correlate positively with productivity. We are not able to measure the selection effects in our data, as exits are uncommon among publicly listed firms and entries are not observed until firms are publicly listed. Nevertheless, we think that tax homogenization would likely further raise productivity by improving the efficiency of firm selection.

[19] More generally, long-run output growth depends on the price elasticities of factor supply. Models with positively-sloped capital supply schedules, such as Aiyagari (1994), would predict smaller gains, and those with positively sloped labor supply schedules would predict larger gains than the figures reported here.

[20] Because $\omega_\tau = \frac{1}{1-\tau}$, the percent change in the tax wedge is related to the percent change in the tax rate as follows: $\frac{d\omega_\tau}{\omega_\tau} = -\frac{d\tau}{\tau} \times \frac{\tau}{1-\tau}$. Multiplying that with the cost elasticity of capital demand, $1/(1-\alpha)$ gives the total effect on capital stock.

4.3 TFP gains in alternative models of labor distortions

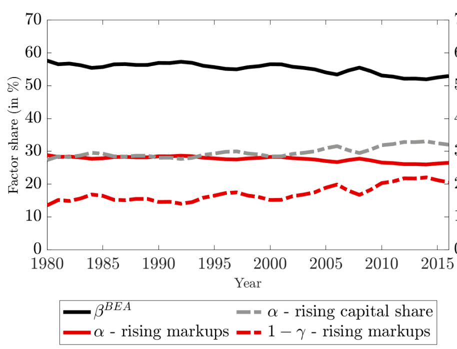
We now turn to alternative models discussed in Section 2.2, where the correlation between labor productivity and corporate tax rates arises structurally. Because that correlation is positive, eliminating tax heterogeneity reduces distortions to employment and improves aggregate TFP over and above what we have estimated so far. The resulting estimate for the TFP gains associated with eliminating tax heterogeneity is shown as the dashed red line in Figure 2. Overall, the estimated gains are considerably larger, averaging 9.1 percent over our sample period. Recall that the average TFP gain in our baseline scenario was 3.3 percent. Gains are particularly large earlier in the sample, averaging around 10 percent. This was a time when the average corporate tax rate was higher. In recent years, the TFP gain stands around 7 percent, suggesting a 3 percent gain from improvements in the distribution of tax asymmetry across firms.

4.4 Macroeconomic trends and estimated TFP gains

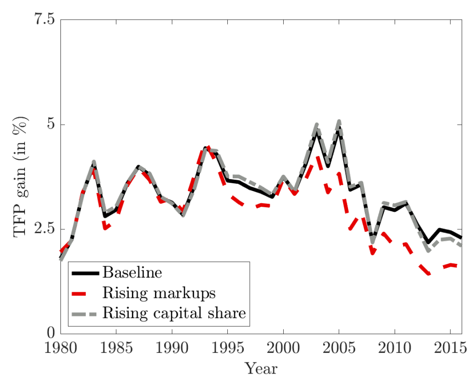
The TFP gains in Figure 2 appear to be roughly stable over time. When computing these gains, we assumed constant values for the macroeconomic parameters of factor shares and markups. Empirically, however, we observe a downward trend in the labor share of income during our sample. Recent work has argued that the decline in the labor share is associated with a rise in price markups and/or an increase in the capital share. In this subsection, we investigate how these changes might affect the trends in estimated TFP gains associated with tax heterogeneity.

We take the decline of the labor share as given and benchmark its drop to the BEA's measure for each year of our analysis. Of course, the shares of capital, labor and profits sum to one, and a change in the labor share necessarily changes at least one other parameter. While the evidence on the decline in the aggregate labor share is relatively well-accepted, how much of that decline was redistributed to capital versus profits is less clear. We therefore consider two alternative scenarios. First, we assume that the decline in the labor share was matched one for one by a rise in the capital share of income, keeping the share of profits constant over time. Second, we assume that the labor share and capital declined proportionally, thus raising the profit share $1-\gamma$ over time. It turns out that neither of these two scenarios changes our findings very much.

The resulting values for α, β, and γ are shown in Panel (a) of Figure 4. The black solid line shows the labor share of income published by the BEA. It declines from 57.7 percent in 1980 to 53.0 percent in 2016. In the first scenario we consider, this is matched by a 4.7 percent increase in the capital share α, shown as the gray dashed line in the figure. In the second scenario, we keep the cost shares of labor and capital fixed at 2/3 and 1/3, implying to a rise in the profit share from 13.5 in 1980 to 20.5 in 2016, or, equivalently, a rise in the markup rate from 15.6 to 25.7 percent (shown as the red dashed line). The resulting capital share decline from 28.8 percent to 26.5 percent is shown as the red solid line.



(a) Parameters



(b) Misallocation

Figure 4: Macro trend scenarios

Note.– Panel (a) shows the labor share of income (β) from the BEA, and the associated changes in the shares of capital (α) and profits ($1-\gamma$) under two alternative scenarios. Panel (b) shows the TFP gains from tax rate equalization in each scenario. Parameters are constant in the baseline scenario.

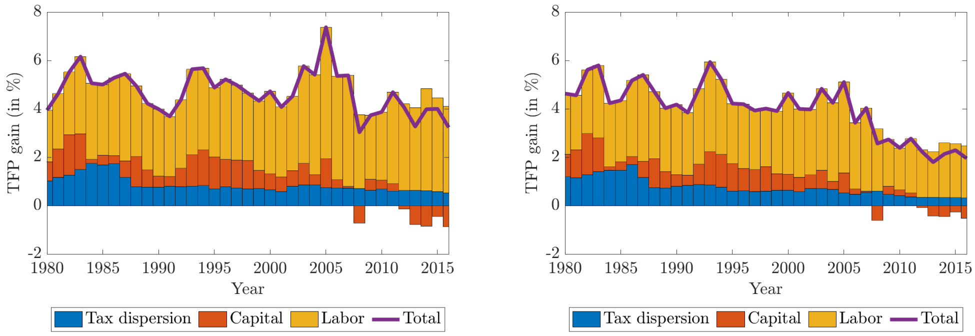
Because our experiments change multiple parameters at once and because the TFP equation in (10) is non-linear, it is *a priori* not clear how the macroeconomic trends might change our estimates. However, we can obtain some insights by studying its approximation in equation (14).

Changes in the labor share of income have a ceteris paribus ambiguous effect on estimated TFP gains. Note that the labor share of income β acts as a weight in equation (14) when considering distortions that are related to the allocation of labor versus capital. A lower β shifts the weight from the correlation between tax distortions and other labor distortions to capital. Because the effect of tax heterogeneity on allocative efficiency is generally ambiguous in a distorted economy - it depends on the correlations of tax rates with other distortions to capital and labor - the net effect of a change in the labor share is ambiguous as well. For instance, if labor productivity is generally high and capital productivity is generally low among firms that face higher tax rates, then a lower labor share should be associated with smaller TFP losses from tax heterogeneity.

We first consider the scenario of a higher capital share. A higher value of α raises the TFP gains, ceteris paribus. This effect is unambiguous from (14). Intuitively, the larger the importance of the distorted factor in production is, the larger are the losses from tax distortions. However, the dashed gray line in panel (b) of Figure 4 shows that the TFP gains from this scenario yield similar magnitudes as our baseline, indicating that the role of the capital share is quantitatively small.

We now consider the second scenario, under which markups increase, which is equivalent to a lower value of γ. Ceteris paribus, lower values of γ are associated with lower TFP losses from tax heterogeneity as indicated by equation (14). Therefore rising markups would tend to reduce the TFP gains over time. Intuitively, this is because lower values of γ bring the economy further away from a linear technology (or, equivalently, from perfect competition), where the best firm can absorb all the resources without facing diminishing returns to scale. The lack of that possibility lowers the total gains to reallocating inputs more efficiently.

Looking at the red dashed line in panel (b) of Figure 4 once again shows



(a) Rising Capital Share

(b) Rising Markups

Figure 5: Sources of TFP gains under alternative macro scenarios

similar TFP gains under this scenario relative to the baseline. Our results point to an approximately 1 percentage point lower potential TFP gain from eliminating tax heterogeneity in recent years.

Decomposition Figure 5 shows the decomposition of TFP gains under alternative macro scenarios where the labor share declines. The left panel attributes that decline to a higher capital share of income and the right panel to a rise in markups. The relative magnitudes of the interaction between tax wedges and capital or labor distortions are broadly similar across scenarios. Under rising markups, overall TFP gains from eliminating tax heterogeneity decline by more, especially toward the end of the sample. This is attributable to a lower value of γ, which reduces the TFP losses associated with each component (see equation (14)), although the decline in the labor component is the most apparent.

5 Conclusion

Our findings show that policies that seek to reduce differences in marginal corporate income tax rates would result in significant gains in aggregate productivity.

The majority of these gains are attributable to the empirical correlation between factor productivity and marginal tax rates: firms that face higher tax rates are typically those where capital and labor are more productive on the margin.

These findings highlight the importance of accounting for other frictions when studying the implications of a specific distortion to aggregate efficiency. In a frictionless economy, or in a distorted economy where tax rates are orthogonal to other economic frictions, the productivity effect of tax asymmetries would have been much smaller in magnitude.

The extent of potential gains in productivity also depends on the model used to interpret the empirical patterns in the data, ranging from around 3 percent in standard models where corporate taxes distort the marginal cost of capital up to 9 percent in models where taxes additionally distort the marginal cost of labor. Empirical assessments of models of production where labor is chosen dynamically, or models with liquidity constraints where payments to labor are made prior to sales, are promising avenues for future research that seeks to ascertain whether and how hiring decisions are causally distorted by corporate tax rates, conditional on capital.

It is also worth noting that although we estimate TFP losses associated with tax heterogeneity, eliminating the heterogeneity altogether is not necessarily productivity maximizing in distorted economies. Tax design that seeks to maximize aggregate productivity would aim to offset other distortions in the economy by strategic variations in the tax rate. Efficient design of corporate taxation is another promising avenue for future research.

References

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia (2022). Misallocation, Selection, and Productivity: A Quantitative Analysis With Panel Data From China. *Econometrica*, 90(3), pp. 1261–1282. doi:https://doi.org/10.3982/ECTA16598.

Aiyagari, S. Rao (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics*, 109(3), pp. 659–684. doi:10.2307/2118417.

Auerbach, Alan J. (1986). The Dynamic Effects of Tax Law Asymmetries. *The Review of Economic Studies*, 53(2), pp. 205–225. doi:10.2307/2297647.

Baqaee, David Rezza and Emmanuel Farhi (2020). Productivity and Misallocation in General Equilibrium. *The Quarterly Journal of Economics*, 135(1), pp. 105–163.

Bils, Mark, Peter J. Klenow, and Cian Ruane (2021). Misallocation or Mismeasurement? *Journal of Monetary Economics*, 124, pp. S39–S56. doi:https://doi.org/10.1016/j.jmoneco.2021.09.004.

Blouin, Jennifer, John E. Core, and Wayne Guay (2010). Have the tax benefits of debt been overestimated? *Journal of Financial Economics*, 98(2), pp. 195–213. doi:https://doi.org/10.1016/j.jfineco.2010.04.005.

Cummins, Jason G., Kevin Hassett, and R. Glenn Hubbard (1994). A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments. *Brookings Papers on Economic Activity*, 25(2), pp. 1–74.

David, Joel M. and Venky Venkateswaran (2019). The Sources of Capital Misallocation. *American Economic Review*, 109(7), pp. 2531–67. doi:10.1257/aer.20180336.

Djankov, Simeon, Tim Ganser, Caralee McLiesh, Rita Ramalho, and Andrei Shleifer (2010). The Effect of Corporate Taxes on Investment and Entrepreneurship. *American Economic Journal: Macroeconomics*, 2(3), pp. 31–64. doi:10.1257/mac.2.3.31.

Gilchrist, Simon, Jae W. Sim, and Egon Zakrajšek (2013). Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs. *Review of Economic Dynamics*, 16(1), pp. 159–176. doi:https://doi.org/10.1016/j.red.2012.11.001.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez (2017). Capital Allocation and Productivity in South Europe. *The Quarterly Journal of Economics*, 132(4), pp. 1915–1967. doi:10.1093/qje/qjx024.

Graham, John R. (1996). Proxies for the corporate marginal tax rate. *Journal of Financial Economics*, 42(2), pp. 187–221. doi:https://doi.org/10.1016/0304-405X(96)00879-3.

Graham, John R. and Lillian F. Mills (2008). Using tax return data to simulate corporate marginal tax rates. *Journal of Accounting and Economics*, 46(2), pp. 366–388. doi:https://doi.org/10.1016/j.jacceco.2007.10.001.

Hall, Robert E. and Dale W. Jorgenson (1967). Tax Policy and Investment Behavior. *The American Economic Review*, 57(3), pp. 391–414.

Harberger, Arnold C. (1962). The Incidence of the Corporation Income Tax. *Journal of Political Economy*, 70(3), pp. 215–240.

Harberger, Arnold C. (1964). The Measurement of Waste. *The American Economic Review*, 54(3), pp. 58–76.

Hopenhayn, Hugo A. (1992). Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica*, 60(5), pp. 1127–1150. doi:10.2307/2951541.

Hsieh, Chang-Tai and Peter J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4), pp. 1403–1448. doi:10.1162/qjec.2009.124.4.1403.

Kaymak, Barış and Immo Schott (2019). Loss-offset provisions in the corporate tax code and misallocation of capital. *Journal of Monetary Economics*, 105, pp. 1–20. doi:10.1016/j.jmoneco.2019.04.011.

Kim, Seho (2023). Optimal Carbon Taxes and Misallocation across Heterogeneous Firm. Technical report, University of Maryland.

Lipsey, R. G. and Kelvin Lancaster (1956). The General Theory of Second Best. *The Review of Economic Studies*, 24(1), pp. 11–32.

Midrigan, Virgiliu and Daniel Yi Xu (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, 104(2), pp. 422–58. doi:10.1257/aer.104.2.422.

Olley, G. Steven and Ariel Pakes (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6), pp. 1263–1297. doi:10.2307/2171831.

Restuccia, Diego and Richard Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4), pp. 707–720. doi:https://doi.org/10.1016/j.red.2008.05.002.

Slattery, Cailin and Owen Zidar (2020). Evaluating State and Local Business Incentives. *Journal of Economic Perspectives*, 34(2), pp. 90–118. doi:10.1257/jep.34.2.90.

A Theoretical Appendix

This appendix details the derivations in the text and gives the formal proofs for the propositions.

Firm's problem with multiple distortions The optimality conditions for investment and future capital of the firm's problem in (1) are given by the first-order condition with respect to investment:

$$i: \quad -1 - \Phi'\left(\frac{i}{k}\right) + q - \mu = 0 \quad \Leftrightarrow \quad q = 1 + \Phi'\left(\frac{i}{k}\right) + \mu. \tag{A1}$$

and

$$k': \quad \beta\left[(1-\tau)zF_k(k',n) + E_{z'|z}V_k(z',k')\right] - q = 0, \tag{A2}$$

with the envelope condition

$$V_k(z,k) = -\Phi\left(\frac{i}{k}\right) + \Phi'\left(\frac{i}{k}\right)\frac{i}{k} + q(1-\delta) + \mu\zeta q. \tag{A3}$$

This implies an optimality condition for the choice of capital given by (5).

Proposition 1. *Total factor productivity in the distorted economy is:*

$$Z = \frac{Y}{K^\alpha N^\beta} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{1-\alpha}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} dG\right]^\beta} \tag{A4}$$

Proof. The profit-maximizing levels of capital, labor, and output are:

$$n = z^{\frac{1}{1-\gamma}} \left(\frac{\beta}{\omega_n}\right)^{\frac{1-\alpha}{1-\gamma}} \left(\frac{\alpha}{\omega_k}\right)^{\frac{\alpha}{1-\gamma}} \tag{A5}$$

$$k = z^{\frac{1}{1-\gamma}} \left(\frac{\beta}{\omega_n}\right)^{\frac{\beta}{1-\gamma}} \left(\frac{\alpha}{\omega_k}\right)^{\frac{1-\beta}{1-\gamma}} \tag{A6}$$

$$y = z^{\frac{1}{1-\gamma}} \left(\frac{\beta}{\omega_n}\right)^{\frac{\beta}{1-\gamma}} \left(\frac{\alpha}{\omega_k}\right)^{\frac{\alpha}{1-\gamma}} \tag{A7}$$

Substituting in the definition of TFP gives the following:

$$Z = \frac{Y}{K^\alpha N^\beta} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{1-\alpha}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} dG\right]^\beta}$$

□

Proposition 2. *Total factor productivity in the counterfactual economy with homogeneous tax rates across firms is:*

$$Z^* = \frac{\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{\beta}{1-\gamma}} \omega_k^{-\frac{1-\beta}{1-\gamma}} \omega_\tau^{\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_n^{-\frac{1-\alpha}{1-\gamma}} \omega_k^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\alpha}{1-\gamma}} dG\right]^\beta}. \tag{A8}$$

Proof. In the counterfactual economy, $\omega_k' = \bar{\omega}_k \omega_R = \bar{\omega}_k \omega_k / \omega_\tau$ and $\omega_n' = \bar{\omega}_n \omega_n$. Proof follows from substituting these in equation (10). The common components of distortions, $\bar{\omega}_n$ and $\bar{\omega}_k$, cancel each other out. □

Proposition 3. *If ω_k, ω_n, and ω_τ are jointly distributed log-normally, then eliminating the heterogeneity in the marginal tax rates yields the following change in aggregate TFP:*

$$\ln \frac{Z^*}{Z} = \frac{1}{2} \frac{1}{1-\gamma} \alpha(1-\beta)\sigma_\tau^2 + \frac{\alpha(1-\beta)}{1-\gamma} \sigma_\tau^2 (L_{k\tau} - 1) + \frac{\alpha\beta}{1-\gamma} \sigma_\tau^2 L_{n\tau}, \tag{A9}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_k$ and $\ln \omega_n$ on $\ln \omega_\tau$.

Proof. Let $\mu_x = \mathbb{E}[\ln x]$ and $\sigma_x^2 = \mathbb{V}[\ln x]$ be the mean and variance of the log of a variable x. Under joint log-normality:

$$
\begin{aligned}
\ln \int y &= \frac{\alpha}{1-\gamma}\ln\alpha + \frac{\beta}{1-\gamma}\ln\beta + \frac{1}{1-\gamma}(\mu_z - \alpha\mu_k - \beta\mu_n) \\
&+ \frac{1}{2}\frac{1}{(1-\gamma)^2}(\sigma_z^2 + \alpha^2\sigma_k^2 + \beta^2\sigma_n^2 - 2\alpha\sigma_{zk} - 2\beta\sigma_{zn} + 2\alpha\beta\sigma_{kn})
\end{aligned}
$$

$$
\begin{aligned}
\ln \int k &= \frac{(1-\beta)}{1-\gamma}\ln\alpha + \frac{\beta}{1-\gamma}\ln\beta + \frac{1}{1-\gamma}(\mu_z - (1-\beta)\mu_k - \beta\mu_n) \\
&+ \frac{1}{2}\frac{1}{(1-\gamma)^2}(\sigma_z^2 + (1-\beta)^2\sigma_k^2 + \beta^2\sigma_n^2 - 2(1-\beta)\sigma_{zk} - 2\beta\sigma_{zn} + 2(1-\beta)\beta\sigma_{kn})
\end{aligned}
$$

$$
\begin{aligned}
\ln \int n &= \frac{\alpha}{1-\gamma}\ln\alpha + \frac{(1-\alpha)}{1-\gamma}\ln\beta + \frac{1}{1-\gamma}(\mu_z - \alpha\mu_k - (1-\alpha)\mu_n) \\
&+ \frac{1}{2}\frac{1}{(1-\gamma)^2}(\sigma_z^2 + \alpha^2\sigma_k^2 + (1-\alpha)^2\sigma_n^2 - 2\alpha\sigma_{zk} - 2(1-\alpha)\sigma_{zn} + 2\alpha(1-\alpha)\sigma_{kn})
\end{aligned}
$$

Using these equations, the TFP in the distorted economy is:

$$
\ln Z = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_k^2 - \beta(1-\alpha)\sigma_n^2 - 2\alpha\beta\sigma_{kn}\right]. \tag{A10}
$$

When tax differentials are eliminated capital distortions are given simply by ω_R, which gives the TFP in that counterfactual as:

$$
\ln Z^* = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_R^2 - \beta(1-\alpha)\sigma_n^2 - 2\alpha\beta\sigma_{Rn}\right]. \tag{A11}
$$

Note that $\sigma_k^2 - \sigma_R^2 = \sigma_\tau^2 + 2\sigma_{R\tau}$ and $\sigma_{kn} - \sigma_{Rn} = \sigma_{n\tau}$. Rearranging the terms and netting out μ_z, σ_z^2 and σ_n^2 terms, the efficiency losses from distortions are equivalent to:

$$
\begin{aligned}
\ln \frac{Z^*}{Z} &= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_k^2-\sigma_R^2)+2\alpha\beta(\sigma_{kn}-\sigma_{Rn})\right] \\
&= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_\tau^2+2\sigma_{R\tau})+2\alpha\beta\sigma_{n\tau}\right] \\
&= \frac{\sigma_\tau^2}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(1+2\frac{\sigma_{R\tau}}{\sigma_\tau^2})+2\alpha\beta\frac{\sigma_{n\tau}}{\sigma_\tau^2}\right] \\
&= \frac{\sigma_\tau^2}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(1+2(L_{k\tau}-1))+2\alpha\beta L_{n\tau}\right] \\
&= \frac{1}{2}\frac{1}{1-\gamma}\alpha(1-\beta)\sigma_\tau^2+\frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1)+\frac{\alpha\beta}{1-\gamma}\sigma_\tau^2 L_{n\tau}
\end{aligned}
$$

The last two equalities substitute the linear projection coefficients for $\frac{\sigma_{R\tau}}{\sigma_\tau^2} = L_{R\tau} = L_{k\tau} - 1$ and $\frac{\sigma_{n\tau}}{\sigma_\tau^2} = L_{n\tau}$. □

A.1 Adding cash-in-advance restrictions to the model

Consider a version of the baseline model in Section (2), where labor costs are incurred before production and sales take place. For comparability, we change our notation slightly to denote the wage costs by ω_l here, and reserve ω_n for the effective marginal cost of labor once taxes and cash-in-advance requirements are taken into account. The firm's optimization problem in this case becomes:

$$
V(z,k) = \max_{n,k',i} \; -i-\omega_l n-\Phi\left(\frac{i}{k}\right)k+\beta\left[(1-\tau)zF(k',n)+\tau\omega_l n+E_{z'|z}V(z',k')\right]
$$

subject to the law of motion for capital and a collateral constraint

$$
\begin{aligned}
k' &= i+(1-\delta)k \\
\omega_l n+i &\leq \zeta q k,
\end{aligned}
$$

with associated multipliers q and μ. Note that wage payments are also subject to the collateral constraint as they need to be financed prior to production. These modifications do not alter the optimality condition for investment. The optimality condition for labor is now given by:

$$zF_n(k,n) = \omega_l + \omega_l \frac{1-\beta+\mu}{\beta} \underbrace{\frac{1}{1-\tau}}_{\omega_\tau} = \omega_n$$

The marginal cost of labor reflects two additional factors in addition to the usual wage cost. Suppose first that the collateral constraint is not binding for the firm: $\mu = 0$. Because there is a lag between the labor outlay and the realization of sales, the cost of labor includes the opportunity cost of paying wages in advance, which is the foregone earnings on those funds that the firm could otherwise enjoy. This is represented by $\omega_l(1-\beta)/\beta$. Because this component is not deducted from corporate income in the optimization problem above, the opportunity cost of paying labor outright, as opposed to investing those funds on the market, interacts with the tax wedge raising the marginal cost of labor by $\omega_n = \omega_\tau \omega_l (1-\beta)/\beta$. This formulation defines a structural relation between the marginal cost of labor ω_n and $\omega_t au$. Higher values of $\omega_t au$ necessarily raise ω_n. Therefore, altering the distribution of $\omega_t au$ also alters the distribution of ω_n. We ignore this possibility in our baseline model.

The second factor originates from the collateral constraint. When the firm is financially constrained ($\mu > 0$) the effective discount factor is higher than $(1-\beta)/\beta$, because the opportunity cost of paying labor in advance is not to earn interest on those funds in the market, but to finance much needed capital on the margin. That the collateral constraint is binding implies that the firm's return on capital is higher than the market. Therefore, constrained firms have a higher marginal cost of labor. Furthermore, tax wedges exacerbate any existing differences in marginal cost of labor that might arise from collateral constraints.

$\sigma_n^2 = \sigma_l^2 + \sigma_x^2 + 2\sigma_{lx}$ where $x = \ln(1 + \frac{1-\beta+\mu}{\beta}\omega_\tau) \simeq \ln\frac{1-\beta+\mu}{\beta} + \ln\omega_\tau)$

Proposition 4. *If $\omega_k = \omega_R \cdot \omega_\tau, \omega_n = \omega_l \cdot \omega_\tau$, and ω_τ are jointly distributed log-*

normally, then eliminating the heterogeneity in the marginal tax rates yields the following change in aggregate TFP:

$$\ln \frac{Z^*}{Z} = \frac{1}{2}\frac{\gamma}{1-\gamma}\sigma_\tau^2 + \frac{\alpha}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1) + \frac{\beta}{1-\gamma}\sigma_\tau^2(L_{n\tau}-1), \tag{A12}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_k$ and $\ln \omega_n$ on $\ln \omega_\tau$.

Proof. Substituting $\omega_n = \omega_l$ and $\omega_k = \omega_R$ in equation (A10) for the TFP in the distorted economy gives:

$$\ln Z = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_k^2 - \beta(1-\alpha)\sigma_n^2 - 2\alpha\beta\sigma_{kn}\right]. \tag{A13}$$

When tax differentials are eliminated capital distortions are given simply by ω_R, which gives the TFP in that counterfactual as:

$$\ln Z^* = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_R^2 - \beta(1-\alpha)\sigma_l^2 - 2\alpha\beta\sigma_{Rl}\right]. \tag{A14}$$

Note that $\sigma_k^2 - \sigma_R^2 = \sigma_\tau^2 + 2\sigma_{R\tau}$, $\sigma_n^2 - \sigma^2 = \sigma_\tau^2 + 2\sigma_{l\tau}$ and $\sigma_{kn} - \sigma_{Rl} = \sigma_{R\tau} + \sigma_{l\tau} + \sigma_\tau^2$. Rearranging the terms and netting out μ_z and σ_z^2 terms, the efficiency gains or losses from distortions are equivalent to:

$$\begin{aligned}
\ln \frac{Z^*}{Z} &= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_k^2 - \sigma_R^2) + \beta(1-\alpha)(\sigma_n^2 - \sigma_l^2) + 2\alpha\beta(\sigma_{kn} - \sigma_{Rl})\right] \\
&= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_\tau^2 + 2\sigma_{R\tau}) + \beta(1-\alpha)(\sigma_\tau^2 + 2\sigma_{l\tau}) \cdots \right. \\
&\qquad \left. \cdots + 2\alpha\beta(\sigma_{R\tau} + \sigma_{l\tau} + \sigma_\tau^2)\right] \\
&= \frac{\sigma_\tau^2}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(1 + 2\frac{\sigma_{R\tau}}{\sigma_\tau^2}) + \beta(1-\alpha)(1 + 2\frac{\sigma_{l\tau}}{\sigma_\tau^2}) \cdots \right. \\
&\qquad \left. \cdots + 2\alpha\beta(1 + \frac{\sigma_{R\tau}}{\sigma_\tau^2} + \frac{\sigma_{l\tau}}{\sigma_\tau^2})\right]. \\
&= \frac{1}{2}\frac{\gamma}{1-\gamma}\sigma_\tau^2 + \frac{\alpha}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1) + \frac{\beta}{1-\gamma}\sigma_\tau^2(L_{n\tau}-1)
\end{aligned}$$

The last equality substitutes the linear projection coefficients for $\frac{\sigma_{R\tau}}{\sigma_\tau^2} = L_{R\tau} = L_{k\tau} - 1$ and $\frac{\sigma_{l\tau}}{\sigma_\tau^2} = L_{l\tau} = L_{n\tau} - 1$, and $\gamma = \alpha + \beta$. □

Proposition 5. *Assume that the tax wedge is measured with error, $\omega_\tau^* = \omega_\tau + \epsilon$ with $\mathbb{E}(\epsilon) = 0$. Then, replacing the ω_τ by ω_τ^* in equation (14) underestimates the net TFP gain from eliminating tax heterogeneity:*

$$\ln(Z^*/Z)|\omega_\tau^* = \ln(Z^*/Z)|\omega_\tau - \frac{\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}, \tag{A15}$$

Proof. Let $\sigma_{\tau^*}^2 = \sigma_\tau^2 + \sigma_\epsilon^2$ denote the variance of the measured tax wedge. Define the projection $x = L_{x\tau} \times \ln \omega_\tau + e_x$, where ω_τ is the true tax wedge. The OLS estimate of $L_{x\tau}$ from the projection of x on $\ln \omega_{\tau^*}^2$ is: $\hat{L}_{x\tau}^{OLS} = L_{x\tau} \times \frac{\sigma_\tau^2}{\sigma_{\tau^*}^2}$.

$$\begin{aligned}
\ln(Z^*/Z)|\omega_\tau^* &= \frac{\alpha(1-\beta)}{1-\gamma}\frac{\sigma_{\tau^*}^2}{2} + \frac{\alpha(1-\beta)}{1-\gamma}\frac{\sigma_{\tau^*}^2}{2}(\hat{L}_{k\tau}^{OLS} - 1) + \frac{\alpha\beta}{1-\gamma}\frac{\sigma_{\tau^*}^2}{2}\hat{L}_{n\tau}^{OLS} \\
&= -\frac{\sigma_\tau^2 + \sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma} + \frac{\alpha(1-\beta)}{1-\gamma}\frac{\sigma_\tau^2}{2}L_{k\tau} + \frac{\alpha\beta}{1-\gamma}\frac{\sigma_\tau^2}{2}\hat{L}_{n\tau} \\
&= \ln(Z^*/Z)|\omega_\tau - \frac{\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}.
\end{aligned}$$

□

A.2 Computation of the non-linear gains and measurement error correction

For each $x \in \{\omega_k, \omega_n\ z\}$, define $\ln \hat{x} = b_{x\tau} \times \ln \omega_\tau$, where $b_{x\tau}$ is the OLS estimator of $L_{x\tau}$. We ignore the intercept term in that projection, because it cancels out from the TFP equations as it is common to all firms. Note also that we are using the OLS estimator, not the consistent estimator.[A1]

[A1]Because various biases from using the OLS estimator cancel each other out, the resulting formula for error correction is simpler and aligns with that reported in Proposition 6. We also considered using the consistent (IV) estimator, and a non-parametric estimator for $\mathbb{E}[\ln x|\omega_t au]$. Both yield similar results.

Then compute current TFP by substituting $\hat{\omega}_k$, $\hat{\omega}_n$ and $\hat{z}$ in equation (10), and the ideal TFP by substituting $\hat{\omega}_n$, $\hat{z}$, and $\omega'_k = \hat{\omega}_k/\omega_{\tau^*}$ for ω_k in the same equation. This yields the following equations:

$$\hat{Z}_{\tau^*} = \frac{\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-\alpha b_{k\tau})} dG_{\tau^*}}{\left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-(1-\beta) b_{k\tau})} dG_{\tau^*}\right]^{\alpha} \left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-(1-\alpha) b_{n\tau}-\alpha b_{k\tau})} dG_{\tau^*}\right]^{\beta}}$$

$$\hat{Z}^*_{\tau^*} = \frac{\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-\alpha b_{R\tau})} dG_{\tau^*}}{\left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-(1-\beta) b_{R\tau})} dG_{\tau^*}\right]^{\alpha} \left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-(1-\alpha) b_{n\tau}-\alpha b_{R\tau})} dG_{\tau^*}\right]^{\beta}},$$

where $b_{R\tau} = b_{k\tau} - 1$ is the projection coefficient of $\ln \omega_R$ on $\ln \omega_{\tau^*}$ and G_{τ^*} is the marginal distribution of the measured tax wedge.

When ϵ is distributed independently log-normal, then for any scalar $c > 0$, $E[\omega^c_{\tau^*}] = E[\omega^c_\tau \epsilon^c] = E[\omega^c_\tau] \cdot E[\epsilon^c] = E[\omega^c_\tau] \cdot \exp(c^2\sigma^2_\epsilon/2)$. Replacing c with the appropriate power component for each term gives:

$$\ln \frac{\hat{Z}^*_{\tau^*}}{\hat{Z}_{\tau^*}} = \ln \frac{\hat{Z}^*_\tau}{\hat{Z}_\tau} - \frac{\sigma^2_\epsilon}{2} \frac{\alpha(1-\beta)}{1-\gamma}. \tag{A16}$$

Repeating these steps for the alternative formulation in Section 2.2 by substituting τ^* for τ in the expression for $\hat{Z}^{**}_{\tau^*}$ presented in equation (16) gives the following relation between the measured TFP loss and the actual TFP loss:

$$\ln \frac{\hat{Z}^{**}_{\tau^*}}{\hat{Z}_{\tau^*}} = \ln \frac{\hat{Z}^{**}_\tau}{\hat{Z}_\tau} - \frac{\sigma^2_\epsilon}{2} \frac{\gamma}{1-\gamma}. \tag{A17}$$

B Data Appendix

The firm-level data used in Section 3 were constructed as follows. We use the annual Compustat database, which provides balance-sheet data on publicly listed companies in the US. Our sample includes the years 1980–2021. We perform the following sample selection and data-cleaning steps. We restrict attention to firms registered in the US. We exclude firms in the finance, insurance, and real estate sectors, as well as in utilities and public administration. We remove observations with negative sales.

We construct firm-level capital stocks by using a perpetual inventory method. For each firm, we start with the year in which information on gross and net property, plant, and equipment (PPEGT and PPENT) is available. We then build the capital stock by adding the change in PPENT deflated by the investment price deflator to the calculated capital stock for that year.

We supplement these data with information on firms' marginal tax rates, taken from two sources, *i)* Graham and Mills (2008), abbreviated as "GM" and *ii)* Blouin et al. (2010), abbreviated as "BCG".[A2] While the GM database covers the years 1980–2021, the BCG data are only available from 1980–2016. We link the Compustat data to the marginal tax rate data via the firm identifier GVKEY. Finally, we remove firm-year observations for which both tax rates are missing.

This results in a sample of 185,203 unique firm-year observations, averaging about 4,600 firms per year. Marginal effective tax rates are available for 70.3 percent of our observations (90.2 percent for the BCG tax rates).

Estimation of firm-level TFP We estimate firm-level TFP using a three-step control function approach following Olley and Pakes (1996). The key variables are value added, employment, and physical capital for each firm and factor shares in the production function. Value added is defined as sales plus the change in inventories.

We begin by estimating factor shares at the two-digit NAICS level in three steps. First, we regress log of output on second-order polynomials in the logs of the capital stock and investment expenditures, including an interaction term as well as log employment. We control for a full set of indicator variables for year and 2-digit NAICS classifications. Sectors with fewer than 100 observations were dropped from this estimation. Second, to correct for survival bias, we estimate a

[A2]Each study contains two measures of the marginal tax rates: before and after interest deductions are applied. The marginal tax rates we use in this paper are after interest deductions.

probit specification for firm survival in the Compustat data (using the same polynomials and year-industry dummies). In a third step, we estimate capital shares for each industry by regressing log output on the log capital stock, controlling for industry-year effects and the predicted survival probability from the previous step. We then compute log TFP assuming a Cobb-Douglas production function and normalize its mean to zero in each year and industry.