Risk Markups

Sebastian Di Tella — Stanford GSB & NBER
Cedomir Malgieri — Arizona State University
Christopher Tonetti — Stanford GSB & NBER

October 29, 2025

We study optimal policy when heterogeneous markups reflect compensation for risk instead of market power. The optimal gross labor tax equals the product of (1) the inverse of the aggregate markup and (2) workers' Pareto weight divided by their consumption share. Markups correctly capture the private cost of risk that reduces labor demand, but they are socially inefficient. This requires a subsidy to undo the aggregate markup, as in the traditional market-power perspective. However, even with lump-sum transfers, uninsurable risk leads to inefficient precautionary saving. Entrepreneurs dynamically overaccumulate too large a share of wealth to self insure and an income effect makes relatively impoverished workers oversupply labor. In the long run this effect dominates and it's optimal to tax labor.

Keywords: Entrepreneurs, Incomplete Risk Sharing, Markups, Wealth Distribution, Optimal Policy

JEL Codes: D52, E21, E23, G11, H21, L26

*We have benefited from discussions with Manuel Amador, Javier Bianchi, Anmol Bhandari, Corina Boar, Joel David, Hugo Hopenhayn, Oleg Itskhoki, Chad Jones, Pete Klenow, Virgiliu Midrigan, Martin Schneider, Yongs Shin, and Gianluca Violante. We thank seminar participants at many universities, conferences, and research centers for useful comments.

1 Introduction

Markups are measured to be large and heterogeneous across firms, which can lower TFP, distort aggregate employment, and create inequality. The role of policy depends on the source of markups. A common view is that markups arise because firms have market power and restrict supply to increase prices and profits. From this perspective, Baqaee and Farhi (2020) and Edmond, Midrigan, and Xu (2023) show that it is optimal to increase output with a subsidy to undo the markup distortions.

However, since at least Knight (1921) economists have recognized that profits may reflect compensation for risk. Providing quantitative support for this view, Boar, Gorea, and Midrigan (2025) attribute the vast majority of the dispersion in firm profits to uninsurable idiosyncratic risk, with a more limited role for market power.

Our main contribution is to study optimal policy when measured markups are compensation for risk. We consider a planner who can use a uniform labor tax and lump-sum transfers. The constrained-efficient allocation implements the first-best level of employment with a gross labor tax equal to the product of (1) the inverse aggregate markup and (2) workers' Pareto weight divided by their consumption share,

$$1 + \tau_{\ell t} = \frac{1}{1 + \pi_t} \times \frac{\gamma_{wt}}{\eta_{wt}}.$$

The first component calls for a labor subsidy to eliminate the aggregate markup, just as in the market-power case. Although markups correctly capture the private cost of risk, they are socially inefficient due to a risk-sharing externality.

Risk markups, however, also distort the consumption distribution relative to the first best. Even though the planner can use lump-sum transfers to achieve any desired redistribution, precautionary saving leads entrepreneurs to dynamically overaccumulate wealth to self insure. Thus, relatively impoverished workers have a consumption share below their Pareto weight and supply too much labor. The consumption-share component of the optimal tax captures this effect. In the long run, the income effect on labor supply dominates the inefficient aggregate markup and the optimal policy is a tax to reduce output, in sharp contrast to the common wisdom derived from the market-power perspective.

To better understand the inefficiency we also consider a planner who can use type-specific labor and savings taxes to control the consumption distribution and address misallocation. The planner uses a modified inverse Euler equation and subsidizes more productive entrepreneurs. However, the formula for the optimal average labor tax is unchanged, and entrepreneurs' consumption share is still larger than their Pareto weight. In this sense, our baseline results are robust.

Finally, we can represent the optimal long-run tax and the welfare gain from implementing any tax rate in terms of a sufficient statistic: workers' consumption as a share of their labor income. We use this sufficient statistic to perform a back-of-the-envelope welfare calculation across a range of values for the aggregate markup and workers' consumption share of income. If measured markups are risk markups, there are modest welfare gains from implementing the optimal tax and large welfare losses from implementing a subsidy equal to the inverse of the aggregate markup.

1.1 Relationship to Existing Literature

Measured markups are large and heterogeneous across firms and can lower TFP (misallocation).[1] A common view is that markups arise because firms have market power. From this perspective, Baqaee and Farhi (2020) and Edmond, Midrigan, and Xu (2023) show that it is optimal to increase output with a subsidy to undo the markup distortions. The logic is simple; since monopolists restrict quantity to increase price, the decentralized equilibrium features lower output than the first best. There are nuances to optimal policy when measured markups reflect market power, such as when there are socially valuable spillovers and firm profits finance fixed costs of innovation, as studied by Cavenaile, Çelik, and Tian (2023) and Adhami, Brouillette, and Rockall (2024). Nonetheless, optimal policy often involves

[1]See Hall (1988), De Loecker and Warzynski (2012), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Midrigan and Xu (2013).

subsidies to producers, especially large producers.

Alternatively, the economy may be efficient if measured markups reflect adjustment costs as in Asker, Collard-Wexler, and De Loecker (2014), search frictions as in Menzio (2024), or if firms are able to perfectly price discriminate as Bornstein and Peter (2024) show. Beyond efficiency concerns, there may be a social preference to reduce inequality. In an economy where wealthy people disproportionately earn capital income, Boar and Midrigan (2025) find that the improvements to efficiency from subsidizing high-markup firms raises wages enough to offset the negative implications of increased capital income on inequality. In contrast to this literature, we study optimal policy when measured markups reflect compensation for risk. Our analysis highlights the role for policy to redistribute in order to address an inefficient wealth distribution, independent of a social preference for equality.

David, Schmid, and Zeke (2022) also study risk and misallocation. In their model different firms have different exogenous loadings on an aggregate risk factor. As in CAPM, the different risk profiles generate different risk premia, which generate deviations from equalizing MPKs, even though the decentralized equilibrium is efficient. In our model risk premia are endogenous, generated by idiosyncratic risk and incomplete markets. This leads to a substantively different perspective on optimal policy. David and Zeke (2024) further explore the implications of aggregate risk and misallocation for optimal monetary policy in a New Keynesian model and David, Ranciere, and Zeke (2023) explore the implications for international diversification and the labor share.

We share a focus on uninsurable idiosyncratic risk with the closely related paper of Boar, Gorea, and Midrigan (2025). Their paper is a quantitative paper that explains the source of heterogeneity in firm returns. In addition to risk, their model features decreasing returns to scale production technology (equivalently monopolistic competition) and collateral constraints. They find that uninsurable idiosyncratic risk is by far the most important feature that explains the empirical variation in returns. Motivated by their quantitative results on the importance of uninsurable idiosyncratic risk, we complement their paper by providing a theoretical analysis of optimal policy in such a world. Since optimal policy is relatively well understood when misallocation is generated by market power or by collateral constraints, we only model uninsurable idiosyncratic risk. This allows us to provide insights from analytical results in a tractable model.

More broadly, we build on the tradition of modeling entrepreneurship as in Evans and Jovanovic (1989), Quadrini (2000), and Cagetti and Nardi (2006). This is relevant for private firms, that are a large part of the economy and are often undiversified (Moskowitz and Vissing-Jørgensen (2002)), but also for many large public firms where ownership is often highly concentrated (e.g., Meta, Amazon, Walmart, and Ford). Our model of entrepreneurship is most similar to that in Moll (2014) and our analysis of optimal policy is similar to Itskhoki and Moll (2019). These papers, however, focus on collateral constraints that prevent firms from operating at their optimal scale, which induces capital misallocation. A key insight from their analysis is that the degree of persistence in the stochastic idiosyncratic productivity process affects the magnitude of steady-state TFP losses from misallocation and the speed of transition towards steady state. Buera and Shin (2013) similarly focus on the effects of collateral constraints on misallocation, in a model with decreasing returns to scale production and an occupational choice, with a focus on quantitative results. To clearly distinguish the new mechanism of risky production in our paper, we make labor the only factor of production and omit collateral constraints.

Uninsurable risky production has also been studied by Meh and Quadrini (2006), Angeletos (2007), and Di Tella and Hall (2022). Because these papers feature iid shocks there is no potential for misallocation and heterogeneous markups across persistently different types. Additionally, Meh and Quadrini (2006) and Angeletos (2007) do not study optimal policy, which is our main contribution. Additionally, Di Tella and Hall (2022) only studies business cycles, whereas we focus on transition dynamics and the long run, leading to insights on the inefficiency of the wealth distribution.

Eeckhout and Veldkamp (2023) is another paper exploring how risk affects our interpretation of measured markups. In their paper, firms acquiring data can reduce risk markups by reducing risk (improving forecasting) but can increase markups by inducing investment to capture market share and exploit market power. Dou, Ji, Tian, and Wang (2025) also studies how risk and incomplete markets affects misallocation, with a focus on how misallocation introduces medium-run fluctuations in TFP that affect aggregate growth and asset prices.

2 The Model

Overview. There is a representative worker and a continuum of entrepreneurs. The workers supply labor, which is the only factor of production, and face no risk. Each entrepreneur is endowed with a distinct linear technology to produce the homogeneous consumption good. There are four essential ingredients. First, producing is risky; even though labor is freely adjusted each instant, entrepreneurs must hire labor before the realization of their productivity shock. Second, entrepreneurs are risk averse. Third, there is incomplete risk sharing; entrepreneurs must retain an equity share in their firm. Fourth, entrepreneurs are heterogeneous in their productivity, risk, and ability to diversify. This heterogeneity introduces the possibility of misallocation. These ingredients deliver heterogeneous risk premia that appear as a wedge in the first order condition for hiring labor. Since the risk premium is the difference between the price of the good and the unit labor cost, we call it the risk markup.

Agents and Preferences. There is a representative worker and a continuum of entrepreneurs, indexed by $i \in [0, 1]$, with preferences

$$U_w(c_w, \ell) = \mathbb{E}\left[\int_0^\infty e^{-\rho_w t}\left(\log c_{wt} - \frac{\ell_t^{1+1/\psi}}{1+1/\psi}\right)dt\right], \tag{1}$$

$$U_e(c_i) = \mathbb{E}\left[\int_0^\infty e^{-\rho_e t}\log c_{it}dt\right]. \tag{2}$$

Technology. Each entrepreneur has a linear technology that uses labor to produce the homogeneous consumption good:

$$dY_{it} = \ell_{it}\ (z_{it}dt + \sigma_{yit}dB_{yit}). \tag{3}$$

where B_{yi} is a Brownian motion specific to the entrepreneur. A central assumption in our setting is that production is risky. At time t the entrepreneur chooses labor ℓ_{it} and its marginal product is $z_{it}dt + \sigma_{yit}dB_{yit}$. The first term is expected productivity and the second is the risk associated with hiring labor to produce.

Each entrepreneur has a multidimensional idiosyncratic state $x_{it} \in \mathbb{R}^N$ that determines his expected productivity, $z_{it} = z(x_{it})$, and his risk, $\sigma_{yit} = \sigma_y(x_{it})$, both bounded, $z(x) \in [0, \bar{z}]$ and $\sigma_y(z) \in [0, \bar{\sigma}_y]$. The idiosyncratic state follows a diffusion

$$dx_{it} = \mu_x(x_{it})dt + \sigma_x(x_{it})dB_{xit}, \tag{4}$$

where B_{xi} is a Brownian motion independent of B_{yi}. We can summarize the behavior of x_{it} with generator $\mathcal{L}_x$. We assume that it converges to a unique non-degenerate stationary distribution $p(x)$, and that the economy starts there.

Resource Constraints. The resource constraints are[2]

$$c_t = \int_0^1 c_{it}di + c_{wt} = \int_0^1 z_{it}\ell_{it}di, \tag{5}$$

$$\ell_t = \int_0^1 \ell_{it}di. \tag{6}$$

Markets and Financial Assets. There are competitive markets for the consumption good, labor, a risk-free bond in zero net supply, and entrepreneurs' equity. The consumption good is the numeraire, with price normalized to 1. Markets are incomplete: entrepreneurs must retain a fraction $\phi_{it} \in [0, 1]$ of their profits, which can also depend on their idiosyncratic

[2]Expressions of the form $\int a_{it}di$ are short-hand for $\int a dG_{at}$ where G_{at} is the cross-sectional measure of variable a at time t. In equation (5) we are assuming a law of large numbers applies. That is, we focus on allocations such that $\int_0^1 \left(\int_0^t \ell_{is}\sigma_y dB_{yis}\right) di = 0$.

state, $\phi_{it} = \phi(x_{it})$. Budget constraints are

$$dn_{wt} = (r_t n_{wt} - c_{wt} + w_t \ell_t)dt, \tag{7}$$

$$dn_{it} = (r_t n_{it} - c_{it} + (z_{it} - w_t)\ell_{it})dt + (\phi_{it}\sigma_{yit})\ell_{it} dB_{yit}. \tag{8}$$

where n_{it} and n_{wt} are agents' bond positions, r_t is the interest rate, and w_t the wage. Entrepreneurs' inability to share idiosyncratic risk is reflected in their budget constraint: they must retain an exposure ϕ_{it} to the idiosyncratic risk in their production.[3] Both entrepreneurs and workers face natural borrowing constraints. For entrepreneurs, this means $n_{it} \geq 0$. Workers cannot borrow more than the present value of labor income, $n_{wt} \geq -\int_t^\infty e^{-\int_t^s r_u du} w_s \ell_s ds$. Let n_{et} be the aggregate net worth of entrepreneurs. Bond market clearing requires $n_{wt} + n_{et} = 0$.

Entrepreneurs may also be able to trade arrow securities contingent on the realization of B_{xit} risk, but since it only affects their investment opportunities, under log preferences they have no hedging motive and will choose not to trade them. For this reason it is without loss of generality to abstract from them (see Appendix A).

Equilibrium Definition. The *representative worker's problem* is to choose $(c_{wt} > 0, \ell_t \geq 0)$ to maximize his utility subject to his budget constraint and natural borrowing constraint, taking prices (r, w) and initial bonds n_{w0} as given. An *entrepreneur's problem* is to choose $(c_{it} > 0, \ell_{it} \geq 0)$ to maximize his utility subject to his budget constraint and natural borrowing constraint, taking prices (r, w) and initial bonds n_{i0} as given. Markets clear if equations (5) and (6) hold (the bond market clears by Walras' Law). A *Competitive Equilibrium* is an allocation (c_w, ℓ_w) and (c_i, ℓ_i) for $i \in [0, 1]$, and market prices (r, w) such that the representative worker and each entrepreneur solve their problem and markets clear.

[3]Entrepreneurs sell, in a competitive market, a claim to a fraction $(1 - \phi_{it})$ of their profits $dY_{it} - w_t \ell_{it} dt$ in exchange for payment $a_{it} dt$. Since the idiosyncratic risk can be diversified, the market prices this claim so $\mathbb{E}\left[(1 - \phi_{it})(dY_{it} - w_t \ell_{it} dt) - a_{it} dt\right] = 0$. Thus, the entrepreneur's budget constraint is $dn_{it} = (r_t n_{it} - c_{it})dt + \phi_{it}(dY_{it} - w_t \ell_{it} dt) + a_{it} dt = (r_t n_{it} - c_{it} + (z_{it} - w_t)\ell_{it})dt + (\phi_{it}\sigma_{yit})\ell_{it} dB_{yit}$.

Discussion of Heterogeneity. The idiosyncratic state x_{it} allows us to model entrepreneur heterogeneity in a general way. A simple example widely used in the firm dynamics literature is $z(x_{it}) = x_{it}$ which follows an AR(1) process, and $\sigma_y(x_{it}) = \sigma_y$ homogeneous and constant. But our environment with multidimensional x allows us to capture richer firm dynamics and permanent types. For example, we can model startups with high productivity growth, high risk, and little diversification that can eventually transition to more diversified corporations with lower growth and lower risk, highlighted by Luttmer (2011). We can also model small undiversified businesses like the plumbers discussed in Hurst and Pugsley (2011) that have stable productivity and never transition to startups or large corporations. This model also permits a corporate sector containing fully diversified but lower productivity firms as in Angeletos (2007).

2.1 Characterizing the Competitive Equilibrium

Optimality Conditions and Risk Markups. The representative worker's optimal allocation is characterized by an Euler equation and the first order condition for labor supply:

$$\frac{dc_{wt}}{c_{wt}} = (r_t - \rho_w)dt, \tag{9}$$

$$c_{wt}^{-1} w_t = \ell_t^{1/\psi}. \tag{10}$$

An entrepreneur's optimal allocation can be characterized by an Euler equation and the first order condition for labor demand:[4]

$$\frac{dc_{it}}{c_{it}} = (r_t - \rho_e + \sigma_{cit}^2)dt + \sigma_{cit} dB_{yit}, \tag{11}$$

$$z_{it} - w_t \leq \sigma_{cit}\phi_{it}\sigma_{yit} \quad \text{(with equality if } z_{it} \geq w_t), \tag{12}$$

where $\sigma_{cit} = \frac{\ell_{it}}{n_{it}}(\phi_{it}\sigma_{yit}) = \frac{\ell_{it}}{c_{it}}(\rho_e\phi_{it}\sigma_{yit})$ is the volatility of consumption growth of entrepreneur i and is at the center of our analysis. Here we have used the fact that with log preferences $c_{it} = \rho_e n_{it}$. The entrepreneur's consumption is therefore locally exposed only

[4] See Appendices A and B for details.

to the idiosyncratic risk in his output B_{yit} (and not the idiosyncratic risk in his idiosyncratic state B_{xit}). The σ_{cit}^2 term in the Euler equation captures the precautionary saving motive.

Entrepreneurs with $z_{it} > w_t$ hire labor ($\ell_{it} > 0$), while less-productive entrepreneurs do not produce. Hiring labor is like investing in a risky asset. The $\sigma_{cit}\phi_{it}\sigma_{yit} \geq 0$ term in equation (12) is a risk premium on labor: it captures the covariation between the marginal utility of the entrepreneur and the fraction of the risk in the marginal product of labor that cannot be shared. Due to log preferences, there is no hedging motive, so the risk premium does not depend on σ_x.

Taking wealth as given, the demand for labor trades off the expected marginal profit from hiring labor, $z_{it} - w_t$, against increased exposure to risk $(\phi_{it}\sigma_{yit})^2$,

$$\ell_{it} = \frac{(z_{it} - w_t)^+}{(\phi_{it}\sigma_{yit})^2} n_{it}. \tag{13}$$

Wealth provides self insurance, so entrepreneurs with more wealth hire more labor. Given wages, entrepreneurs hire less labor compared to the complete markets case. Even though production of the homogeneous good is linear, the most productive entrepreneurs do not hire all the labor because that would expose them to too much risk. Homothetic preferences and linear budget constraints result in labor demand that, when positive, is linearly increasing in bonds n_{it} (or equivalently, consumption c_{it}), and decreasing in uninsurable risk $\phi_{it}\sigma_{yit}$.

The risk premium on labor creates labor wedges at the entrepreneur level that look like markups. We define the risk markup, $1 + \pi_{it}$, to be the expectation of price over unit cost. For entrepreneurs who produce,

$$1 + \pi_{it} := \frac{1}{w_t/z_{it}} = 1 + \frac{\sigma_{cit}\phi_{it}\sigma_{yit}}{w_t} > 1. \tag{14}$$

Markups depend only on productivity z_{it} relative to the market wage w_t. Large markup firms are high-productivity firms. Even highly diversified firms can have a large markup. Diversification (low ϕ_{it}) shows up as a large size (high ℓ_{it}, and thus high σ_{cit}), not as small markups.

Although the model generates heterogeneous measured markups, the risk-adjusted marginal product of labor, $z_{it} - \sigma_{cit}\phi_{it}\sigma_{yit}$, is equalized to the labor cost, w_t, for all entrepreneurs who have positive labor demand. Thus, there is no true markup in the sense that the properly-measured marginal cost is equal to the price of the good for all active entrepreneurs.

Aggregation and Separation Property. At the aggregate level, risk markups show up as an aggregate markup and as misallocation that reduces TFP,

$$Z_t := \frac{c_t}{\ell_t} = \int_0^1 z_{it}\frac{\ell_{it}}{\ell_t}di, \tag{15}$$

$$1 + \pi_t := \frac{1}{w_t/Z_t} = \int_0^1 \frac{1}{w_t/z_{it}}\frac{\ell_{it}}{\ell_t}di. \tag{16}$$

Aggregate productivity, Z_t, is the labor-weighted average idiosyncratic productivity and the aggregate risk markup, $1 + \pi_t$, is the labor-weighted average idiosyncratic risk markup. Risk markups reduce TFP, because they decrease the labor demand of high productivity entrepreneurs compared to complete markets (where $Z_t = w_t = \bar{z}$).

The environment with log preferences and linear production allows for easy aggregation and a powerful separation result. We can solve for the cross-sectional allocations separately from aggregate employment and output. This property will be essential when we study efficiency in the next section. Without it, the planner problem would quickly become intractable.

First, instead of keeping track of net worth n_{it}, we can equivalently work with $\eta_{it} := c_{it}/c_t$, the consumption share of entrepreneur i. Let $g_t(\eta, x)$ be the joint density of η and x. Because labor demand is linear in η_{it}, when computing equilibrium allocations we do not need to keep track of the joint density g. Define $\eta_t(x)$ as the aggregate consumption share of entrepreneurs of type x, η_{et} the consumption share of all entrepreneurs, and η_{wt} the consumption share of the representative worker.

$$\eta_t(x) := \int \eta g_t(\eta, x)dx, \quad \eta_{et} := \int \eta_t(x)dx, \quad \eta_{wt} := 1 - \eta_{et}. \tag{17}$$

Second, we combine labor demand from equation (13) with the resource constraints to obtain expressions for the wage and aggregate productivity as functions of the distribution $\eta_t(x)$.

That is, we implicitly define the functions $w(\eta_t)$ and $Z(\eta_t)$ by

$$\int z(x)\eta_t(x)\frac{1}{\rho_e}\frac{(z(x)-w)^+}{(\phi(x)\sigma_y(x))^2}dx = 1, \tag{18}$$

$$\left(\int \eta_t(x)\frac{1}{\rho_e}\frac{(z(x)-w)^+}{(\phi(x)\sigma_y(x))^2}dx\right)^{-1} = \frac{c_t}{\ell_t} = Z_t. \tag{19}$$

An important property follows from this. Entrepreneurs' exposure to risk depends only on their type x and the distribution $\eta_t(x)$,

$$\sigma_c(x;\eta_t) = \frac{(z(x)-w(\eta_t))^+}{\phi(x)\sigma_y(x)}. \tag{20}$$

Specifically, it does not depend on aggregate employment or output. This property will play a key role in the efficiency analysis. Together with the Euler equations, it implies that the evolution of the joint distribution $g_t(x,\eta)$ is independent of aggregate labor ℓ_t.[5] Since σ_{cit} only depends on $\eta_t(x)$, we can obtain a simple Kolmogorov forward equation for $\eta_t(x)$,

$$\partial_t\eta_t(x) = \mathcal{L}_x^*\eta_t(x) + A(x;\eta_t), \tag{21}$$

where $\mathcal{L}_x^*$ is the adjoint of $\mathcal{L}_x$ and $A(x;\eta_t)$ is the geometric drift of η_{it} times $\eta_t(x)$.[6] The key result is that this is a self-contained differential equation for $\eta_t(x)$ that characterizes its evolution and does not depend on aggregate labor ℓ_t.

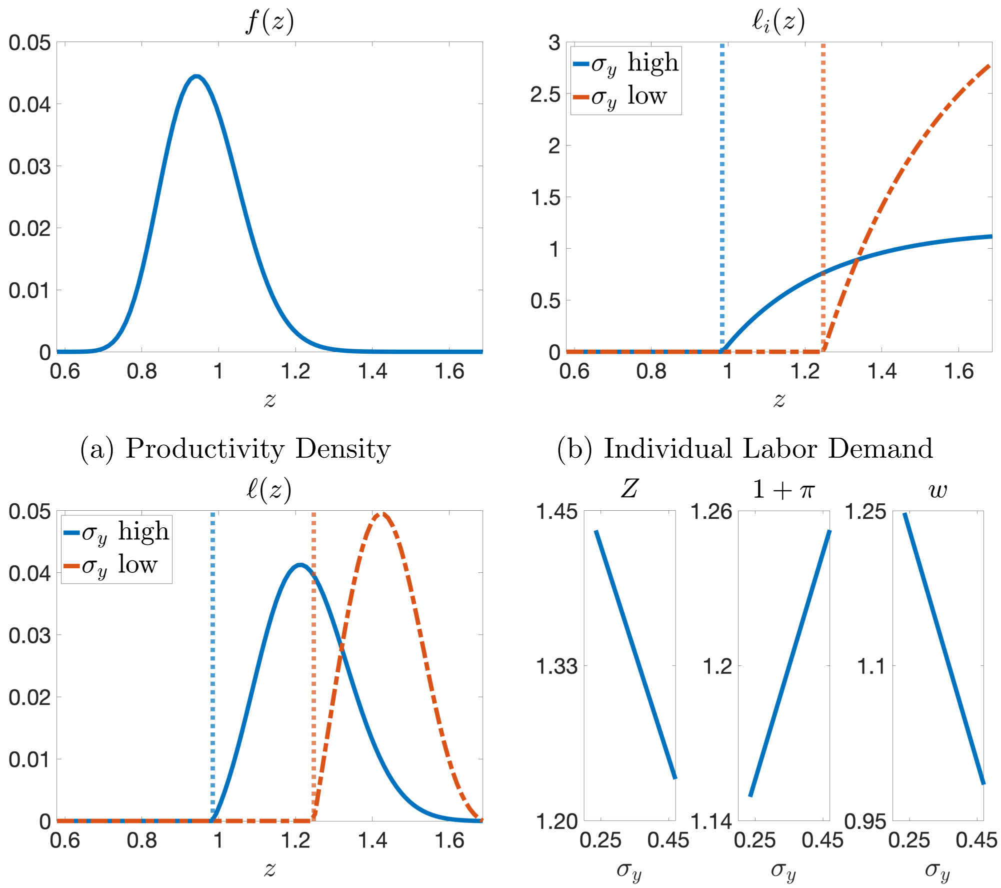
This separation result allows us to compute an equilibrium in two steps. In step one, start with a given initial distribution $\eta_0(x)$. Then, solve the KFE (21) forward to obtain the path of $\eta_t(x)$ and use equations (18)–(19) to solve for the path of w_t and Z_t. Due to the separation result, step one does not involve aggregate output c_t, aggregate employment ℓ_t, or interest rates r_t. Finally, using $c_{wt} = \eta_{wt}Z_t\ell_t$, we use the representative worker's labor supply condition to solve for aggregate employment ℓ_t, the definition of aggregate TFP to compute aggregate output c_t, and the worker's Euler equation to obtain the interest rate r_t. It is then straightforward to construct c_{it}, ℓ_{it}, and n_{it} (see Appendix C for details).

[5] Consumption shares η_{it} follow $d\eta_{it} = \eta_{it}\mu_\eta(x_{it};\eta_t) + \eta_{it}\sigma_c(x_{it};\eta_t)dB_{yit}$, where $\mu_\eta(x;\eta_t) = (\eta_{wt}\rho_w + \eta_{et}\rho_e) - \rho_e + \sigma_c^2(x;\eta_t) - \int \eta_t(x)\sigma_c^2(x;\eta_t)dx$ and the generator is $\mathcal{L}_\eta(x;\eta_t)$. The joint generator for (η,x) is $\mathcal{L}_{x\eta}(\eta_t) = \mathcal{L}_x + \mathcal{L}_\eta(x;\eta_t)$ and the KFE for g is $\partial_t g_t(x,y) = \mathcal{L}_x^* + \mathcal{L}_\eta^*(x;\eta_t)$, independent of ℓ_t.

[6] See Appendix G for details of the KFE derivation.

The separation between the consumption distribution and aggregate labor has major implications for efficiency and optimal policy, as we will explore in Section 3.

2.2 A Numerical Example: Equilibrium



(a) Productivity Density (b) Individual Labor Demand

(c) Total Labor Hired by Type-z (d) TFP, Markup, and Wage

Figure 1: Comparative Static for σ_y

Before turning to social efficiency, we illustrate key characteristics of the competitive equilibrium through analysis of comparative statics in a numerical example. This provides a visualization of the analytic results presented in the previous section. For simplicity, we set $x_{it} = z_{it}$ with log productivity following an Ornstein-Uhlenbeck process, $\sigma_{yit} = z_{it}\sigma_y$, and $\phi_{it} = 1$. See Appendix H for details. Figure 1 explores how long-run labor demand, TFP, and the aggregate markup change when hiring labor becomes more risky (increase in

σ_y). Blue solid lines are from the baseline calibration and orange dotted lines are from a counterfactual equilibrium with a smaller σ_y. The vertical lines mark equilibrium wages.

Figure 1a plots the stationary productivity density. Figure 1b shows that an entrepreneur's labor demand is positive when productivity is higher than the wage and increases in productivity (holding fixed wealth). When production is riskier, entrepreneurs do not scale their firm size with their productivity as aggressively. Figure 1c plots the distribution of total labor hired by all entrepreneurs of type z. Three forces determine the distribution of labor across productivity types. First is the density of z. Second, the entrepreneur's optimal labor hiring policy is increasing in productivity. Third, labor demand is a function of the bond distribution.

In this numerical example, there is not much labor hired by low productivity types close to the wage threshold even though there may be many entrepreneurs of this type because they each hire little labor. Total labor hired by z-types initially increases in z even when the density of entrepreneurs is decreasing, because each entrepreneur is hiring more workers. Finally, as z further increases, the density of entrepreneurs falls fast enough that the total labor hired by z-type entrepreneurs falls, even as each entrepreneur continues to hire more workers.

Figure 1d shows the effect of increased risk on key macro aggregates. In an economy where production is riskier, each entrepreneur would hire less labor holding the wage constant. In equilibrium, the wage falls, generating entry of previously inactive entrepreneurs into production. The end result is that the largest producers reduce labor demand and the smallest producers increase labor demand. The aggregate markup rises, reflecting the increase in risk premia. In this model risk generates misallocation, lowering TFP. Connecting to the misallocation literature, more risk (higher σ_y) or less risk sharing (higher ϕ) could be a potential explanation for differences in cross-country TFP levels and the size distribution of firms. We now turn to analyzing the efficiency properties of the competitive equilibrium.

3 Efficiency

We consider two approaches to constrained efficiency. An organizing principle is that the planner cannot complete the market. We study a baseline case which is policy-relevant and facilitates comparison between the risk markup and market power perspectives. We also study a case in which we saturate the planner with a rich set of policy instruments, even though we do not think they are realistic, in order to better understand the source of inefficiency in the equilibrium.

Baseline Case. In our baseline the planner has access to a uniform labor tax and lump-sum transfers across agents at time $t = 0$.[7] Motivated by asymmetric information concerns, we study uniform taxes that do not depend on idiosyncratic characteristics.[8] The $t = 0$ transfers can accomplish any redistribution, so taxes are used only for efficiency purposes. We also consider restrictions on the planner's ability to redistribute.

Saturated-Instrument Case. In the saturated-instrument case, the planner has enough instruments to control all margins but cannot complete the market. In this case, the planner can target taxes by entrepreneur type and can also control the wealth distribution with taxes on savings. A key result is that the average labor tax formula is the same as in the baseline case. In this sense our baseline results are robust and this section helps us to understand them.

The optimal labor tax has a component that is the deviation of the workers' consumption share from the first best. In the baseline case, the planner cannot control the wealth distribution, so this deviation arises from entrepreneurs precautionary saving. This raises the question of why the planner doesn't just use an entrepreneur-specific tax on savings to undo their precautionary saving. In the saturated-instrument case, we give the planner the ability

[7] A tax on labor is equivalent to a production tax. A tax on realized profits would complete the market. A uniform savings tax on entrepreneurs and workers is equivalent to lump-sum transfers.

[8] For example, firms can merge on paper to harvest size-dependent subsidies without changing any economic activity, or can misreport their activity (Chen, Liu, Serrato, and Xu, 2021).

to use such taxes to control the wealth distribution. Even then it is not constrained efficient to set consumption shares to first best and the labor tax formula is unchanged with a component that captures the deviation in consumption shares from first best. A key takeaway from both cases is that the consumption distribution and savings margin plays an important role for efficiency when measured markups are compensation for risk.

3.1 Planner's objective function

The Planner's objective function is

$$\int_0^\infty e^{-\rho_w t}\tilde{\Gamma}_w\left(\log c_{wt} - \frac{\ell_t^{1+1/\psi}}{1+1/\psi}\right)dt + \int_0^1 \mathbb{E}\left[\int_0^\infty e^{-\rho_e t}\tilde{\Gamma}_e \log c_{it}dt\right]di, \tag{22}$$

where $\tilde{\Gamma}_w > 0$ and $\tilde{\Gamma}_e > 0$ are the Pareto weights on the representative worker and entrepreneurs. For simplicity of exposition, we assume that the planner does not care about entrepreneurs individually (i.e., $\tilde{\Gamma}_i = \tilde{\Gamma}_e \ \forall\, i$). It will be useful to work with discounted Pareto weights and relative discounted Pareto weights,

$$\Gamma_{et} = e^{-\rho_e t}\tilde{\Gamma}_e, \quad \Gamma_{wt} = e^{-\rho_w t}\tilde{\Gamma}_w, \quad \gamma_{et} = \frac{\Gamma_{et}}{\Gamma_{wt}+\Gamma_{et}}, \quad \gamma_{wt} = \frac{\Gamma_{wt}}{\Gamma_{wt}+\Gamma_{et}}. \tag{23}$$

The objective function can be written to separate aggregate labor ℓ_t and the distribution of consumption shares η_{it}:[9]

$$\begin{aligned}&\overbrace{\int_0^\infty (\Gamma_{wt}+\Gamma_{et})\log \ell_t - \Gamma_{wt}\frac{\ell_t^{1+1/\psi}}{1+1/\psi}dt}^{\text{labor}} + \overbrace{\int_0^\infty (\Gamma_{wt}+\Gamma_{et})\log Z_t dt}^{\text{TFP}} \\ &+ \underbrace{\int_0^\infty \Gamma_{wt}\log \eta_{wt} + \Gamma_{et}\left(\int p(x)\log\frac{\eta_t(x)}{p(x)}dx\right)dt}_{\text{workers vs. entrepreneurs}} - \underbrace{\int_0^\infty \Gamma_{et}\left(\int\left(p(x)\log\frac{\eta_t(x)}{p(x)} - \int \log \eta g_t(\eta,x)d\eta\right)dx\right)dt}_{\text{entrepreneur risk}}.\end{aligned} \tag{24}$$

Aggregate labor only appears in the first term. The second term involves TFP. The third term captures the distribution of consumption between workers and entrepreneurs of each type ($\eta_t(x)/p(x)$ is the per-capita consumption of entrepreneurs of type x). The fourth term

[9] Write $c_{wt} = \eta_{wt}Z_t\ell_t$ and $c_{it} = \eta_{it}Z_t\ell_t$, apply logs, and use a law of large numbers to eliminate the expectation. Then add and subtract $\int_0^\infty \Gamma_{et}\left(\int p(x)\log\frac{\eta_t(x)}{p(x)}dx\right)dt$ and regroup terms.

reflects that the planner dislikes inequality among entrepreneurs because they are risk averse. Within entrepreneurs of type x, $\int \log \eta g_t(\eta, x) d\eta$ is their aggregate utility, while $p(x) \log \frac{\eta_t(x)}{p(x)}$ would be their aggregate utility if they all had the average consumption of that type. The fourth term is weakly positive due to Jensen's inequality.

3.2 First Best

The first-best allocation optimizes each term in the objective function. (1) Aggregate labor is set to maximize the first term. (2) Labor is employed only by the highest productivity types. (3) Consumption is split between entrepreneurs and workers according to Pareto weights. (4) There is perfect risk sharing and all entrepreneurs have the same consumption, so the last term is zero. The first-best allocation is

$$\ell_t^{FB} = \gamma_{wt}^{-\frac{\psi}{1+\psi}}, \tag{25}$$

$$Z_t^{FB} = \bar{z}, \tag{26}$$

$$\eta_{wt}^{FB} = \gamma_{wt}, \tag{27}$$

$$\eta_{it}^{FB} = \frac{\eta_t^{FB}(x)}{p(x)} = \eta_{et}^{FB} = \gamma_{et}. \tag{28}$$

Implementation of the First-Best Allocation. With complete markets the competitive equilibrium achieves the first best and the planner implements it with only redistribution at $t = 0$ and no taxes.[10] Agents perfectly share risk, $\sigma_{cit} = 0$, so there are no risk markups, $w_t = \bar{z}$, and only the most productive entrepreneurs produce, $Z_t = \bar{z}$. Workers and entrepreneurs follow their Euler equations without precautionary saving or risk, so $\eta_{it} = \eta_{et}$, and η_{et} follows

$$d\eta_{et} = \eta_{et}(1 - \eta_{et})(\rho_e - \rho_w)dt. \tag{29}$$

This yields $\eta_{et} = \frac{e^{-\rho_e t}\eta_{e0}}{e^{-\rho_w t}\eta_{w0} + e^{-\rho_e t}\eta_{e0}}$. If the planner chooses the initial distribution $\eta_{e0} = \gamma_{e0}$, we obtain $\eta_{et} = \gamma_{et}$. Aggregate labor would then be given by equation (10), which given w_t, Z_t,

[10] This is an application of the first and second welfare theorems.

and η_{wt}, yields the first-best path for aggregate labor.

3.3 Baseline Case: Uniform Labor Tax

The planner has access to a uniform labor tax $\tau_{\ell t}$ paid by entrepreneurs and rebated lump-sum to the workers. We consider two options for the planner's ability to redistribute: (1) he has access to transfers at time $t = 0$ to achieve any desired redistribution, or (2) he must take the initial distribution $\eta_0(x)$ as given (for example, if agents can insure against planner interventions). In both cases, the labor tax is used only for efficiency purposes.

Let $w_t = (1+\tau_{\ell t})\tilde{w}_t$ be the after-tax wage paid by entrepreneurs, and $\tilde{w}_t$ the wage workers earn. The tax rebate is $T_t = \tilde{w}_t \ell_t \tau_{\ell t}$. The only change in the equilibrium conditions is that the worker's labor supply uses $\tilde{w}_t$,

$$\ell_t^{1/\psi} = c_{wt}^{-1} \tilde{w}_t = (\eta_{wt} Z_t \ell_t)^{-1} \frac{w_t}{1+\tau_{\ell t}}. \tag{30}$$

The labor tax allows the planner to control aggregate labor ℓ_t,

$$\ell_t = \left[\eta_{wt} \; (1+\pi_t) \; (1+\tau_{\ell t})\right]^{-\frac{\psi}{1+\psi}}. \tag{31}$$

Because of the separation property of the competitive equilibrium, the labor tax does not affect the distribution $\eta_t(x)$ or any function that depends on it such as $\pi(\eta_t)$. In fact, the entire joint distribution $g(x, \eta)$ is unaffected by ℓ_t.

We can therefore recast the planner's problem as choosing a path for aggregate labor ℓ_t to maximize the first term of equation (24), and choosing an initial distribution $\eta_0(x)$ to maximize the remaining terms of the planner's objective.[11] A property of this formulation is that the choices of ℓ and the distribution of $\eta_0(x)$ can be made independently, which delivers analytical tractability.

A key result is immediate: The planner sets labor to the first-best level. Then the optimal labor tax is obtained by rearranging equation (31).

[11]See Appendix D for details on how to construct an optimal allocation from the solution to this problem.

The Optimal Labor Allocation and Tax

$$\ell_t = \ell_t^{FB} = \gamma_{wt}^{-\frac{\psi}{\psi+1}} \tag{32}$$

$$1 + \tau_{\ell t} = \frac{1}{1+\pi_t} \times \frac{\gamma_{wt}}{\eta_{wt}} \tag{33}$$

The labor tax has a subsidy component that eliminates the effect of the aggregate markup $1 + \pi_t$ on aggregate employment, just as in the market-power setting but for completely different reasons. Although risk markups are compensation for risk, they are nonetheless socially inefficient. However, it also has a component that eliminates the income effect on labor supply caused by distortions in consumption shares relative to the first best ($\eta_{wt}^{FB} = \gamma_{wt}$). Workers supply more labor when their consumption share is lower (see equation (31)). To offset this income effect the planner needs a labor tax if workers consumption share is below the first best or a labor subsidy if the consumption share is above first best.

Understanding the Inefficiency. The competitive equilibrium is inefficient because of incomplete risk sharing,

$$\sigma_{cit} = \frac{\phi_{it}\sigma_{yit}}{n_{it}}\ell_{it} = \rho_e \frac{\phi_{it}\sigma_{yit}}{\eta_{it}} \frac{1}{Z_t} \frac{\ell_{it}}{\ell_t}. \tag{34}$$

Private entrepreneurs take their net worth n_{it} as given, while the planner realizes it is endogenous,

$$n_{it} = \frac{1}{\rho_e} Z_t \ell_t \eta_{it}. \tag{35}$$

When the planner chooses aggregate employment, ℓ_t, he understands that he affects aggregate output $Z_t\ell_t$. The key insight is that higher output increases entrepreneurs' wealth, and since wealth is insurance against risk, this improves their risk sharing. For fixed consumption shares, η_{it}, higher output $c_t = Z_t\ell_t$ raises entrepreneurs' net worth.[12] In the background, the past interest rates increase, according to the Euler equations, and entrepreneurs enter

[12]With complete markets this would be a pecuniary externality without welfare effects. However, since entrepreneurs use their bonds to self insure against uninsurable idiosyncratic risk, it has welfare implications.

period t with higher n_{it}.

To be concrete, equation (34) shows that if an entrepreneur increases his labor demand ℓ_{it}, he increases his exposure to risk, for which he requires compensation (this is where risk markups come from). But if all entrepreneurs raise employment, this raises all entrepreneurs' net worth proportionally, so they are not bearing more risk (in the sense that σ_{cit}^2 is unchanged). This force leads entrepreneurs to hire too little labor. Although risk markups correctly measure the private cost of risk exposure, this is not the social cost of risk. For this reason, the first term in the uniform labor tax eliminates the aggregate markup.

The second term corrects the effect on labor supply caused by entrepreneurs' precautionary saving. If workers' consumption share was as in the first best, $\eta_{wt} = \eta_{wt}^{FB} = \gamma_{wt}$, then the optimal labor tax would equal the inverse of the aggregate markup. However, in contrast to aggregate labor, the consumption distribution is distorted relative to the first best, $\eta_t(x) \neq \eta_t^{FB}(x)$, because of entrepreneurs' precautionary saving. Without precautionary saving entrepreneurs consumption share would satisfy the first-best law of motion in equation (29).[13] Since the planner cannot eliminate precautionary saving, he cannot implement the first-best consumption distribution, even though he can use transfers to control η_0. However, he can still implement the first-best labor allocation by undoing the income effect on labor supply. This is what the second term in the uniform labor tax accomplishes.

Since the consumption distribution is central to our result, in Section 3.4 we study the saturated-instrument planner problem when he can tax savings and control the consumption distribution. A key result is that the labor tax formula is the same and the consumption distribution is still not first best. In this sense, these results are robust.

It is Optimal to Tax Labor, not Subsidize. We can obtain sharper results about the optimal labor tax in the long run and, with some assumptions, on the whole path. First, we can

[13]Use entrepreneurs' Euler equations with the σ_{ct}^2 term eliminated in the drift and aggregate.

manipulate the expression for the optimal uniform labor tax to obtain

$$1 + \tau_{\ell t} = \zeta_t \times \gamma_{wt}, \tag{36}$$

where

$$\zeta_t = \frac{w_t \ell_t}{c_{wt}} \tag{37}$$

is the total labor income, $w_t \ell_t = \tilde{w}_t \ell_t + T_t$, relative to worker's consumption c_{wt}.[14]

Assume that in the long-run the distribution $\eta_t(x)$ converges to some stationary distribution $\eta^*(x)$. This coincides with the long-run distribution of the competitive equilibrium and requires $\rho_e > \rho_w > 0$. This implies $\pi_t = \pi(\eta_t) \to \pi^*$, $w_t(\eta_t) \to w^*$, $\eta_{wt} \to \eta_w^*$, and $\zeta_t \to \zeta^*$. Since $\rho_e > \rho_w$, the optimal tax converges to

$$1 + \tau_\ell^* = \frac{1}{1 + \pi^*} \times \frac{1}{\eta_w^*}, \tag{38}$$

or equivalently,

The Optimal Long-run Labor Tax is Positive

$$1 + \tau_\ell^* = \zeta^* > 1 \tag{39}$$

Using the separation result that consumption shares and markups are independent of the labor tax, ζ^* is independent of planner control. Entrepreneur wealth must be positive, $n_e^* > 0$, so by market clearing $n_w^* < 0$, and thus the budget constraint of the worker implies $c_w^* < w^* \ell^*$. In the long run, since entrepreneurs have accumulated positive wealth, their consumption is larger than their profits. Since aggregate consumption equals aggregate profits plus total labor income (including transfers), workers must therefore consume less than their total labor income. Thus we obtain a key result. In the long-run, the optimal labor tax is positive.

These insights carry over when $\rho_e = \rho_w$. Even though there is no stationary wealth distribution, in the long run workers' consumption share is still less than their Pareto weight

[14] Replace $\frac{1}{1+\pi_t} = \frac{w_t}{Z_t}$ and multiply above and below by ℓ_t and use $c_{wt} = Z_t \ell_t \eta_{wt}$.

and it is still optimal to tax, not subsidize, labor (see Appendix F).

We can further characterize the optimal path of the tax when the initial distribution starts at the steady state $\eta_0 = \eta^*$. This is the optimal tax if the economy started at steady state and the planner could not redistribute at time zero. The optimal tax simplifies to

$$1 + \tau_{\ell t} = (1 + \tau_\ell^*) \times \gamma_{wt}. \tag{40}$$

When the Planner does not value entrepreneurs welfare ($\tilde{\Gamma}_e \to 0$), we have $\gamma_{wt} \to 1$ and therefore $\tau_{\ell t} = \tau_\ell^* > 0$ for all t. When $\tilde{\Gamma}_e > 0$, the optimal tax $\tau_{\ell t}$ is increasing over time. As long as the planner does not care too much about entrepreneurs, if $\tilde{\Gamma}_e < \tau_\ell^* \times \tilde{\Gamma}_w$, then it is optimal to impose a tax along the entire transition path, i.e., $\tau_{\ell t} > 0$ for all time.

Quantitative Exploration of Welfare Gains. In the $\rho_e > \rho_w$ case with a stationary wealth distribution in the competitive equilibrium, we can use our sufficient statistic long-run results to compute the welfare gains across steady states of different tax policies relative to the zero-tax competitive equilibrium. It is useful to rewrite labor supply using equation (31) as

$$\ell_{\tau_\ell}^* = \left(\frac{1 + \tau_\ell^*}{\zeta^*} \right)^{-\frac{\psi}{1+\psi}}, \tag{41}$$

recalling that ζ^*, long-run total labor income divided by workers' consumption, is invariant to the labor tax. In the long run, the only term in the planner's objective that is affected by the labor tax is the first term, so we can compute welfare differences across labor tax policies by isolating the effect on the first term:

$$\begin{aligned} \upsilon(\tau_\ell^*) &:= \log \ell_\tau^* - \frac{\psi}{1+\psi} (\ell_\tau^*)^{\frac{1+\psi}{\psi}} \\ &= -\frac{\psi}{1+\psi} \left(\log[1 + \tau_\ell^*] - \log \zeta^* + \frac{\zeta^*}{1 + \tau_\ell^*} \right). \end{aligned} \tag{42}$$

We let $\Delta\upsilon(\tau_\ell^*)$ be the consumption equivalent welfare gain of tax τ_ℓ^*, defined as the proportional increase in consumption in the zero-tax equilibrium that would make the worker indifferent. Then,

$$
\begin{aligned}
\log\left[1+\Delta\upsilon(\tau_\ell^*)\right] &= \upsilon(\tau_\ell^*)-\upsilon(0) \\
&= -\frac{\psi}{1+\psi}\left(\log\left[1+\tau_\ell^*\right]-\zeta^*\frac{\tau_\ell^*}{1+\tau_\ell^*}\right). \qquad (43)
\end{aligned}
$$

While the reader can input their preferred estimates for these objects, we will report a range of welfare losses associated with benchmark values. We assume either a 20% or 30% average markup, consistent with the range reported in Edmond, Midrigan, and Xu (2023). We consider a Frisch elasticity of 1 or 3 consistent with common values in the macroeconomics literature. For the fraction of labor income that workers consume, we are not aware of a good dataset that allows classification of individuals into workers vs. entrepreneurs and measures their consumption and income. Thus we explore results for a range of values: $c_w = 0.85w\ell$, $c_w = 0.9w\ell$, $c_w = 0.95w\ell$.

We first consider the optimal tax $1+\tau_\ell^* = \zeta^*$, which yields first best labor. The welfare gain is

$$
\log\left[1+\Delta\upsilon(\zeta^*)\right] = -\frac{\psi}{1+\psi}\left(\log\zeta^*-(\zeta^*-1)\right). \qquad (44)
$$

The welfare gain does not depend on the markup. Across all values considered, welfare gains range from 0.07% to 1.35%.

In contrast, we compute the welfare change a planner would achieve if he set the labor tax to exactly undo the markup, as would be done by a planner in this economy with the mistaken belief that markups reflect market power. The welfare change is

$$
\log\left[1+\Delta\upsilon\left(\frac{1}{1+\pi^*}\right)\right] = -\frac{\psi}{1+\psi}\left(-\log\left[1+\pi^*\right]+\zeta^*\pi^*\right). \qquad (45)
$$

Across all values considered, the welfare gain is negative and ranges from -6.57% to -1.98%. The largest welfare gain from the optimal policy is less than the smallest welfare costs of the market-power subsidy policy. In this sense, the welfare loss caused by setting a tax to undo the markup is large and it may be better to refrain from intervention than to tax at the wrong rate.

We can also compute welfare gains along the transition path, but this requires setting values for all parameters of the model (see Appendix H.3). According to a benchmark calibration, the welfare gain from implementing the dynamic optimal policy is 0.092%, the steady-state to steady-state welfare difference is 0.117%, and the welfare gain along the transition of implementing a constant tax equal to the long-run optimal value is 0.087%. In this sense, the long-run results are a good guide to welfare along the transition.

Summary of Results We can now make a straightforward comparison between the risk markup and market-power perspectives. A central tenet in the market-power perspective is that market power leads to underproduction, and should be corrected with subsidies to undo the effect of the markup (see Baqaee and Farhi (2020) and Edmond, Midrigan, and Xu (2023)). In contrast, when markups are compensation for risk, there is overproduction in the long run, which should be corrected with taxes. The planner wants to implement the first-best level of employment, just as in the market-power case. Risk markups are pushing labor supply down and the planner needs a labor subsidy to counteract them. This is just like the in the market power perspective, but for completely different reasons. However, that is not the planner's only concern. Due to entrepreneurs' precautionary saving, the consumption distribution is distorted relative to the first best, which induces an income effect on workers that the planner needs to counteract. In the long run, workers' consumption share is below the first best so the income effect drives excess employment and the planner needs a labor tax to offset it. A key insight of this paper is that the single feature of uninsurable idiosyncratic risk produces not only markups but also distortions in saving behavior.

In short, although risk markups are compensation for risk, they are nonetheless socially inefficient. Furthermore, the risk markup perspective highlights the centrality of the consumption distribution in optimal policy in the face of measured markups. A back-of-the-envelope welfare calculation suggests the welfare gain from the optimal tax is modest, while the welfare cost of a subsidy that offsets markups is large.

3.4 Understanding the Inefficiency: Saturated-Instrument Planner

In the baseline, the optimal tax depends on the consumption distribution, which is distorted from first best by entrepreneurs' precautionary saving. If precautionary saving is inefficient, why doesn't the planner use entrepreneur-specific taxes on savings to undo it? How would access to savings taxes affect the optimal use of the labor tax? In this section we give the planner the ability to use such taxes to control the consumption distribution. We show that precautionary saving is inefficient, but it is not constrained efficient to set consumption shares to first best. Furthermore, the labor tax formula is unchanged with a component that captures the deviation in consumption shares from first best. The saturated planner problem also allows us to better understand the source of inefficiency and how the planner would use targeted taxes to reduce misallocation.

The planner can use labor and savings taxes conditional on the entrepreneurs' state x_{it}, in addition to lump-sum transfers at time zero. The entrepreneur's budget constraint is

$$dn_{it} = (n_{it}r_t - c_{it} + \ell_{it}(z_{it} - \tilde{w}_t(1 + \tau_{\ell t}(x_{it})))\,dt - n_{it}dT_{nit} + \ell_{it}\sigma_{ity}\phi_{it}dB_{yit}, \tag{46}$$

and the natural borrowing limit is $n_{it} \geq 0$. The tax on labor $\tau_{\ell t}(x)$ can depend on the type x. There is also a tax on savings,

$$dT_{nit} = \tau_{nt}(x_{it})dt + \sigma_{nt}(x_{it})dB_{xit}. \tag{47}$$

The tax on savings can be contingent on the realization of the B_{xit} shock through the $\sigma_{nt}(x_{it})$ term. This corresponds to forcing entrepreneurs to trade Arrow securities contingent on B_{xit} (equivalently taxing or subsidizing those Arrow securities). Such Arrow securities are available in the competitive equilibrium, but entrepreneurs are uninterested in dynamically hedging this risk because of log preferences. We are allowing the planner to force entrepreneurs to dynamically hedge. For simplicity, all rebates are given to workers (this is without loss of generality, since the planner has lump-sum transfers).

With these expanded instruments the planner can control the consumption distribution

$\eta_t(x)$ and the share of aggregate employment by entrepreneurs of type x, $\hat{\ell}(x)$. He can therefore directly control TFP,

$$Z_t = \int z(x)\hat{\ell}_t(x)dx, \tag{48}$$

but he faces the following constraint arising from incomplete risk sharing,

$$\sigma_{ct}(x) = \frac{\rho_e \phi(x)\sigma_y(x)}{Z_t \eta_t(x)}\hat{\ell}_t(x). \tag{49}$$

When the planner makes entrepreneurs hire more labor, he must expose them to more risk. Incomplete risk sharing creates dispersion in consumption across entrepreneurs, so the last term in the planner's objective function (24) cannot be set to zero. Within entrepreneurs of current type x, consumption and labor will be dispersed due to their previous history of B_{yit} and B_{xit} shocks. While he cannot eliminate the dispersion due to B_{yit} shocks (i.e., he can't complete markets), he can and will eliminate the dispersion arising from B_{xit} shocks.

We can rewrite the last term in the planner's objective function (24) to obtain

$$\begin{aligned}&\overbrace{\int_0^\infty (\Gamma_{wt}+\Gamma_{et})\log \ell_t - \Gamma_{wt}\frac{\ell_t^{1+1/\psi}}{1+1/\psi}dt}^{\text{labor}} + \overbrace{\int_0^\infty (\Gamma_{wt}+\Gamma_{et})\log Z_t dt}^{\text{TFP}} \\ &+ \underbrace{\int_0^\infty \Gamma_{wt}\log \eta_{wt} + \Gamma_{et}\left(\int p(x)\log\frac{\eta_t(x)}{p(x)}dx\right)dt}_{\text{workers vs. entrepreneurs}} - \underbrace{\int_0^\infty \Gamma_{et}\frac{1}{\rho_e}\frac{1}{2}\left(\int p(x)\sigma_{ct}^2(x)dx\right)dt}_{\text{entrepreneur risk}}.\end{aligned} \tag{50}$$

The planner's problem is to choose aggregate employment ℓ_t, the consumption distribution $\eta_t(x)$, and allocate labor to entrepreneur types $\hat{\ell}_t(x)$, subject to incomplete risk sharing (49), aggregate TFP (48) and resource constraints $\int \hat{\ell}_t(x)dx = 1$ and $\eta_{wt} = 1 - \int \eta_t(x)dx$. See Appendix E for details.

3.4.1 Optimal allocation

Aggregate employment ℓ_t appears separately in the planner's objective function and does not enter constraints. The planner therefore chooses the first-best level of aggregate employment,

$\ell_t = \ell_t^{FB}$. The distribution of labor is given by,

$$\hat{\ell}_t(x) = \left[z(x) \left(\gamma_{wt} + \gamma_{et} \left(1 + \frac{1}{\rho_e} \sigma_{ct}^2 \right) \right) - Z_t \right]^+ \frac{Z_t \eta_t(x)}{\rho_e \left(\phi(x) \sigma_y(x) \right)^2 \gamma_{et} \frac{p(z)}{\eta_t(x)}}. \tag{51}$$

The planner allocates more labor to entrepreneurs who are more productive, have less uninsurable risk, and are wealthier.

The planner deviates from the first-best consumption distribution in equation (27) to improve risk sharing,

$$\frac{\eta_t(x)}{p(x)} = \frac{\gamma_{et} \left(1 + \frac{1}{\rho_e} \sigma_{ct}^2(x) \right)}{\gamma_{wt} + \gamma_{et} \left(1 + \frac{1}{\rho_e} \sigma_{ct}^2 \right)}, \tag{52}$$

where

$$\sigma_{ct}^2 = \int p(x) \sigma_{ct}^2(x) dx \tag{53}$$

is the average consumption risk. Effectively, entrepreneurs' Pareto weights are increased to $\gamma_{et} \left(1 + \frac{1}{\rho_e} \sigma_{ct}^2(x) \right)$. The reason is that the planner understands that giving more consumption (equivalently, wealth) to entrepreneurs of type x not only delivers utility directly, but also improves their risk sharing. The aggregate consumption share of entrepreneurs is

$$\eta_{et} = \frac{\gamma_{et} \left(1 + \frac{1}{\rho_e} \sigma_{ct}^2 \right)}{\gamma_{wt} + \gamma_{et} \left(1 + \frac{1}{\rho_e} \sigma_{ct}^2 \right)}. \tag{54}$$

The consumption share of entrepreneurs is always above first best $\eta_{et} > \eta_{et}^{FB} = \gamma_{et}$, so the consumption share of workers is always below first best.

In summary, due to the separation property the planner implements the first-best level of employment, but he distorts the consumption and labor distributions relative to the first best to improve risk sharing.

3.4.2 Optimal Labor Taxes

The same reasoning as in the baseline case with only a uniform labor tax yields the optimal average labor tax, $1+\tau_{\ell t} := \int (1+\tau_{\ell t}(x))\hat{\ell}_t(x)dx$,

$$1+\tau_{\ell t} = \frac{1}{1+\pi_t} \times \frac{\gamma_{wt}}{\eta_{wt}}. \tag{55}$$

The aggregate markup is

$$1+\pi_t = \frac{Z_t}{\int \tilde{w}_t(1+\tau_{\ell t}(x))\hat{\ell}_t(x)dx} = \int \frac{z(x)}{\tilde{w}_t(1+\tau_{\ell t}(x))} \frac{\hat{\ell}(x)\tilde{w}_t(1+\tau_{\ell t}(x))}{\int \hat{\ell}(x)\tilde{w}_t(1+\tau_{\ell t}(x))dx} dx,$$

which can be measured directly at the aggregate level (aggregate TFP over aggregate after-tax labor cost) or at the entrepreneur level (entrepreneur productivity over after-tax labor cost) and then aggregated with cost weighting.

The average optimal tax has the same formula as in the baseline uniform labor tax case. In this sense the optimal labor tax is robust to richer policy instruments. The values of the average markup and the consumption distribution depend on the policy instruments available to the planner. Nevertheless, however the average markup and consumption distribution are generated, the formula for the average labor tax is unchanged. Furthermore, because workers' consumption share is always below their Pareto weight, the optimal gross labor tax is always larger than the inverse aggregate markup, $1+\tau_{\ell t} > \frac{1}{1+\pi_t}$.

Comparing equilibrium conditions with the planner's optimality, we obtain an expression for the cross-section of the labor tax,

$$\tilde{w}_t(1+\tau_{\ell t}(x)) - Z_t = -\gamma_{et}\frac{1}{\rho_e}\sigma_{ct}^2 z(x) + \sigma_{ct}(x)\phi(x)\sigma_y(x)\left[\gamma_{et}\frac{p(x)}{\eta_t(x)} - 1\right]. \tag{56}$$

The planner subsidizes entrepreneurs with high productivity, $z(x)$, and those whose consumption shares are larger than first best.

Understanding the Labor Inefficiency. The competitive equilibrium is still inefficient because of incomplete risk sharing. Now, in addition to controlling aggregate labor demand, the planner can control TFP with type-specific labor taxes. In the baseline case raising employment increased output and aggregate net worth, which the entrepreneurs used to self insure. Increasing TFP has the same effect.

$$n_{it} = \frac{1}{\rho_e} Z_t \ell_t \eta_{it}. \tag{57}$$

High-productivity entrepreneurs don't internalize that when they all hire more labor they increase TFP and therefore improve risk sharing for everyone. This is captured by the first term in the cross-sectional optimal tax formula (56).

Both the average and cross-sectional labor tax depends on the deviation of the consumption distribution from first best. When the planner only has access to a uniform labor tax, the consumption distribution is not first-best due to the precautionary saving of entrepreneurs and the above logic holds. Even when the planner controls the consumption distribution, he chooses not to set it to the first-best in order to improve risk sharing. In both cases, the deviation of the consumption distribution from the first best distorts aggregate labor supply through workers' income effect. The planner undoes this effect using the labor tax. This is captured in the $\frac{\gamma_{wt}}{\eta_{wt}}$ term in the uniform and average labor tax. The distorted consumption distribution also distorts the allocation of labor across entrepreneurs, which causes misallocation and lowers TFP. The planner counteracts misallocation using the second term in the cross-sectional labor tax. If the planner implemented the first best consumption distribution, $\eta_t(x) = \eta_t^{FB}(x)$, these terms would be absent. When the planner cannot control the consumption distribution, these taxes are obviously needed. Furthermore, even when the consumption distribution is controlled, the planner chooses not to implement the first best consumption distribution and these labor taxes are needed. We now study optimal savings taxes to understand why.

3.4.3 Optimal Savings Taxes

Optimal savings taxes are

$$\tau_{nt}(x) = \sigma_{ct}^2(x) - \mathcal{L}_{xt} \frac{\eta_t(x)/p(x)}{\eta_{wt}} \frac{\gamma_{wt}}{\gamma_{et}}, \tag{58}$$

$$\sigma_{nt}(x) = \frac{\partial_x \left[\frac{\eta_t(x)/p(x)}{\eta_{wt}} \frac{\gamma_{wt}}{\gamma_{et}} \right]}{\frac{\eta_t(x)/p(x)}{\eta_{wt}} \frac{\gamma_{wt}}{\gamma_{et}}} \sigma_x(x), \tag{59}$$

where $\mathcal{L}_{xt} = \mathcal{L}_x + \partial_t$ is the joint generator for (x, t) (see Appendix E). The first-term in equation (58) eliminates the precautionary motive, $\sigma_{ct}^2(x)$. The second term in equation (58) and $\sigma_{nt}(x)$ in equation (59) reflect that the planner deviates from the first-best consumption distribution to improve risk sharing. If $\eta_t(x) = \eta_t^{FB}(x)$ they would both be zero.

Understanding the Consumption Distribution Inefficiency. The consumption distribution η_{it} in the competitive equilibrium is inefficient because incomplete risk sharing induces precautionary saving. When entrepreneurs hire labor they face risk, so why shouldn't they save to self insure? Social optimality requires the marginal cost of delivering utility to entrepreneurs to be a martingale. In standard settings, the marginal cost is the inverse marginal utility of consumption, and this yields the inverse Euler equation as an optimality condition. With complete markets, the Euler equation and the inverse Euler equation coincide. With incomplete markets, however, precautionary saving creates a wedge between the Euler and inverse Euler equations. In the case with only uniform labor taxes, it is no surprise that the consumption distribution is not first best. Here, when the wealth distribution is controlled, why isn't it optimal to implement the first-best?

By itself, the first term in the optimal saving tax $\tau_{nt}(x)$ eliminates the precautionary saving motive and would implement the inverse Euler equation and first-best consumption distribution. In our setting, however, there is an extra twist. The marginal cost of delivering utility to entrepreneurs is not the inverse marginal utility of consumption, c_{it}. Giving more consumption (equivalently wealth) to entrepreneurs also improves insurance, so the marginal cost of utility is $c_{it}(1+\sigma_{ct}^2(x_{it})/\rho_e)^{-1}$. This is why the planner deviates from the first-best consumption allocation in equation (52), as if the Pareto weight was $\gamma_{et}\left(1+\frac{1}{\rho_e}\sigma_{ct}^2(x_{it})\right)$. This is implemented by the second term in equation (58) and by equation (59).[15]

The main takeaway is that, whether the planner controls the consumption distribution or not, incomplete risk sharing leads to a consumption distribution that deviates from first best. How the consumption distribution deviates from first best depends on the instruments the planner has, which ultimately has knock-on effects for the labor taxes needed to implement first-best aggregate labor.

4 Conclusion

We build a competitive model in which measured markups are compensation for risk. When entrepreneurs scale up and hire more workers they are exposed to more risk. Markups correctly capture the private cost of risk, but they are socially inefficient. Entrepreneurs have precautionary savings to self insure against risk, which is also socially inefficient. They overaccumulate wealth over time, leading relatively impoverished workers to supply too much labor. In the long run, the income effect on labor supply dominates the inefficient risk premia and the optimal policy is a tax to reduce output, in sharp contrast to the common wisdom derived from the market-power perspective. Using a sufficient statistic representation of the optimal tax, we show that if measured markups are risk markups, there are modest welfare gains from implementing the optimal tax and large welfare losses from implementing a subsidy equal to the inverse of the aggregate markup.

[15] If $\sigma_{ct}^2(x_{it})$ were constant, the planner would implement the inverse Euler equation. In our setting, along the transition path and due to time varying types x_{it}, the inverse Euler equation is modified.

References

Adhami, M., J.-F. Brouillette, and E. Rockall (2024): "Variable Markups, Incomplete Pass-Throughs, and R&D Misallocation," *working paper.*

Angeletos, G.-M. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10(1), 1–30.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic Inputs and (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013–1063.

Baqaee, D. R., and E. Farhi (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135(1), 105–163.

Boar, C., D. Gorea, and V. Midrigan (2025): "The Aggregate Costs of Uninsurable Business Risk," *American Economic Review*, R&R.

Boar, C., and V. Midrigan (2025): "Markups and Inequality," *The Review of Economics Studies*, forthcoming.

Bornstein, G., and A. Peter (2024): "Nonlinear Pricing and Misallocation," *working paper.*

Buera, F. J., and Y. Shin (2013): "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 121(2), 221–272.

Cagetti, M., and M. D. Nardi (2006): "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 114(5), 835–870.

Cavenaile, L., M. A. Çelik, and X. Tian (2023): "Are Markups Too High? Competition, Strategic Innovation, and Industry Dynamics," *working paper.*

Chen, Z., Z. Liu, J. C. S. Serrato, and D. Y. Xu (2021): "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, 111(7), 2065–2100.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101(3), 471–75.

David, J., R. Ranciere, and D. Zeke (2023): "Risk-Taking, Capital Allocation and Monetary Policy," *working paper.*

David, J., L. Schmid, and D. Zeke (2022): "Risk-Adjusted Capital Allocation and Misallocation," *Journal of Financial Economics*, 145(3), 684–705.

David, J., and D. Zeke (2024): "International Diversification, Reallocation, and the Labor Share," *working paper.*

De Loecker, J., and F. Warzynski (2012): "Markups and Firm-Level Export Status," *American Economic Review*, 102(6), 2437–2471.

Di Tella, S., and R. Hall (2022): "Risk Premium Shocks Can Create Inefficient Recessions," *The Review of Economic Studies*, 89(3), 1335–1369.

Dou, W. W., Y. Ji, D. Tian, and P. Wang (2025): "Misallocation and Asset Prices," *working paper.*

Edmond, C., V. Midrigan, and D. Y. Xu (2023): "How Costly Are Markups?," *Journal of Political Economy*, 131(7), 1619–1675.

Eeckhout, J., and L. Veldkamp (2023): "Data and Markups: A Macro-Finance Perspective," *working paper.*

Evans, D. S., and B. Jovanovic (1989): "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, 97(4), 808–827.

Hall, R. E. (1988): "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96(5), 921–47.

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

Hurst, E., and B. Pugsley (2011): "What do Small Businesses Do?," *Brookings Papers on Economic Activity*, 44(2), 73–142.

Itskhoki, O., and B. Moll (2019): "Optimal Development Policies with Financial Frictions," *Econometrica*, 87(1), 139–173.

Knight, F. H. (1921): *Risk, Uncertainty, and Profit.* Houghton Mifflin Company.

Luttmer, E. G. (2011): "On the Mechanics of Firm Growth," *Review of Economic Studies*, 78(3), 1042–1068.

Meh, C. A., and V. Quadrini (2006): "Endogenous Market Incompleteness with Investment Risks," *Journal of Economic Dynamics and Control*, 30(11), 2143–2165.

Menzio, G. (2024): "Markups: A Search-Theoretic Perspective," *working paper*.

Midrigan, V., and D. Xu (2013): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104(2), 422–458.

Moll, B. (2014): "Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation?," *The American Economic Review*, 104(10), 3186–3221.

Moskowitz, T. J., and A. Vissing-Jørgensen (2002): "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?," *American Economic Review*, 92(4), 745–778.

Quadrini, V. (2000): "Entrepreneurship, Saving, and Social Mobility," *Review of Economic Dynamics*, 3(1), 1–40.

Restuccia, D., and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11(4), 707–720.

Salgado, S. (2020): "Technical Change and Entrepreneurship," *working paper*.

A Entrepreneur Optimality

The entrepreneurs' problem is to chose adapted processes $(c_i > 0, \ell_i \geq 0)$ to maximize utility

$$U(c_i) = \mathbb{E}\left[\int_0^\infty e^{-\rho_e t} \log c_{it} dt\right] \tag{A.1}$$

subject to the dynamic budget constraint

$$dn_{it} = (r_t n_{it} - c_{it} + \ell_{it}(z_{it} - w_t))dt + \ell_{it}\sigma_{yit}\phi_{it} dB_{yit} \tag{A.2}$$

with initial bonds n_{i0} given and the natural borrowing limit $n_{it} \geq 0$.

Any plan (c_i, ℓ_i) satisfying the dynamic budget constraint satisfies the intertemporal budget constraint

$$\mathbb{E}\left[\int_0^\infty \xi_{it} c_{it} dt\right] \leq n_{i0}, \tag{A.3}$$

where ξ_{it} is the solution to

$$d\xi_{it}/\xi_{it} = -r_t dt - \frac{(z_{it} - w_t)^+}{\phi_{it}\sigma_{yit}} dB_{yit}, \tag{A.4}$$

with $\xi_{i0} = 1$. To see this, use Itô's lemma to compute

$$\begin{aligned} d(\xi_{it} n_{it}) &= \xi_{it}(r_t n_{it} - c_{it} + \ell_{it}(z_{it} - w_t) - r_t n_{it} - \frac{(z_{it} - w_t)^+}{\phi_{it}\sigma_{yit}} \ell_{it}\phi_{it}\sigma_{yit})dt + \xi_{it}\left(\ell_{it}\phi_{it}\sigma_{yit} - \frac{(z_{it} - w_t)^+}{\phi_{it}\sigma_{yit}} n_{it}\right) dB_{yit}, \\ &= \xi_{it}(-c_{it} + \ell_{it}\left((z_{it} - w_t) - (z_{it} - w_t)^+\right))dt + \xi_{it}\left(\ell_{it}\phi_{it}\sigma_{yit} - \frac{(z_{it} - w_t)^+}{\phi_{it}\sigma_{yit}} n_{it}\right) dB_{yit}. \end{aligned}$$

Integrate and take expectations to obtain

$$\mathbb{E}\left[\xi_{i\tau_j} n_{i\tau_j}\right] = \xi_{i0} n_{i0} + \mathbb{E}\left[\int_0^{\tau_j} \xi_{it}(-c_{it} + \ell_{it}\left((z_{it} - w_t) - (z_{it} - w_t)^+\right))dt\right],$$

where $\{\tau_j\}$ is an increasing sequence of stopping times with $\lim_{j\to\infty} \tau_j = \infty$ a.s., such that

$\int_0^{\tau_j} \xi_{it} \left(\ell_{it}\phi_{it}\sigma_{yit} - \frac{(z_{it}-w_t)^+}{\phi_{it}\sigma_{yit}} n_{it} \right) dB_{yit}$ is a martingale. Reorganizing we get

$$\mathbb{E}\left[\int_0^{\tau_j} \xi_{it}(c_{it} - \ell_{it}\left((z_{it} - w_t) - (z_{it} - w_t)^+\right))dt\right] + \mathbb{E}\left[\xi_{i\tau_j} n_{i\tau_j}\right] = \xi_{i0} n_{i0}.$$

Notice that $\xi_{it}(c_{it} - \ell_{it}((z_{it} - w_t) - (z_{it} - w_t)^+)) \geq \xi_{it} c_{it} \geq 0$. Take limits as $j \to \infty$ using the monotone convergence theorem and recall $\xi_{it} n_{it} \geq 0$ for any plan that satisfies the natural borrowing constraint. We obtain

$$\mathbb{E}\left[\int_0^{\infty} \xi_{it} c_{it} dt\right] \leq n_{i0}.$$

We will therefore consider a relaxed problem: choose consumption c_i to maximize utility $U(c_i)$ subject to the intertemporal budget constraint (A.3) and then later check that the proposed solution satisfies the dynamic budget constraint. We have a first-order condition

$$e^{-\rho_e t} c_{it}^{-1} = \lambda \xi_{it}, \tag{A.5}$$

where the value of $\lambda > 0$ ensures the constraint holds with equality:

$$\mathbb{E}\left[\int_0^{\infty} \xi_{it} c_{it} dt\right] = \mathbb{E}\left[\int_0^{\infty} e^{-\rho_e t} \lambda^{-1} dt\right] = \frac{1}{\rho_e \lambda} = n_{i0},$$
$$\implies \lambda = \frac{1}{\rho_e} \frac{1}{n_{i0}}.$$

These conditions are sufficient for optimality of the relaxed problem. To see this, consider any alternative plan $\tilde{c}_i$ satisfying the intertemporal budget constraint. Use the gradient inequality for concave functions to note that $\log \tilde{c}_{it} \leq \log c_{it} + c_{it}^{-1}(\tilde{c}_{it} - c_{it})$. Then,

$$\begin{aligned}
U(\tilde{c}_i) &= \mathbb{E}\left[\int_0^{\infty} e^{-\rho_e t} \log \tilde{c}_{it} \, dt\right] \\
&\leq \mathbb{E}\left[\int_0^{\infty} e^{-\rho_e t} \left(\log c_{it} + c_{it}^{-1}(\tilde{c}_{it} - c_{it}\right) dt\right] \\
&= U(c_i) + \mathbb{E}\left[\int_0^{\infty} \lambda \xi_{it}(\tilde{c}_{it} - c_{it}) dt\right] \leq U(c_i).
\end{aligned}$$

The final inequality holds because $\mathbb{E}\left[\int_0^{\infty} \xi_{it}\tilde{c}_{it}\right] \leq n_{i0}$ since $\tilde{c}$ satisfies the intertemporal budget constraint in equation (A.3) while $\mathbb{E}\left[\int_0^\infty \xi_{it} c_{it}\right] = n_{i0}$ by construction.

We now want to derive the stochastic processes that these sufficient conditions imply. From the first-order condition (A.5), we obtain that $c_{it} > 0$ is an Itô process,

$$dc_{it}/c_{it} = \mu_{cit} dt + \sigma_{cit} dB_{yit}.$$

Using Itô's lemma we obtain

$$d(e^{-\rho_e t} c_{it}^{-1}) = d(\lambda \xi_{it}) \tag{A.6}$$

$$e^{-\rho_e t} c_{it}^{-1} \left(-\rho_e - \mu_{cit} + \sigma_{cit}^2\right) dt - e^{-\rho_e t} c_{it}^{-1} \sigma_{cit} dB_{yit} = -\lambda \xi_{it} r_t dt - \lambda \xi_{it} \frac{(z_{it} - w_t)^+}{\phi_{it} \sigma_{yit}} dB_{yit}. \tag{A.7}$$

Matching terms we get the Euler equation and the first-order condition for labor,

$$\mu_{cit} = r_t - \rho_e + \sigma_{cit}^2, \tag{A.8}$$

$$\sigma_{cit} = \frac{(z_{it} - w_t)^+}{\phi_{it} \sigma_{yit}}. \tag{A.9}$$

And any plan $c_i > 0$ satisfying these conditions and $c_{i0} = \lambda^{-1} = \rho_e n_{i0}$ satisfies FOC (A.5) and the intertemporal budget constraint with equality and is therefore optimal in the relaxed problem.

It only remains to show that our candidate plan c_i can indeed be implemented as part of a plan (c_i, ℓ_i) that satisfies the dynamic budget constraint. We set

$$\ell_{it} = \frac{\sigma_{cit} c_{it}}{\phi_{it} \sigma_{yit} \rho_e} \geq 0, \tag{A.10}$$

We now define the process for bonds $n_{it} = c_{it}/\rho_e > 0$ which satisfies the initial condition by construction. Using the optimality conditions we compute

$$
\begin{aligned}
dn_{it}/n_{it} &= \left(r_t - \rho_e + \sigma_{cit}^2\right) dt + \sigma_{cit} dB_{yit}, \\
&= \left(r_t - c_{it}/n_{it} + \frac{\ell_{it}}{n_{it}} \phi_{it}\sigma_{yit} \frac{(z_{it} - w_t)^+}{\phi_{it}\sigma_{yit}}\right) dt + \frac{\ell_{it}}{n_{it}} \phi_{it}\sigma_{yit} dB_{yit}, \\
&= \left(r_t - c_{it}/n_{it} + \frac{\ell_{it}}{n_{it}} (z_{it} - w_t)\right) dt + \frac{\ell_{it}}{n_{it}} \phi_{it}\sigma_{yit} dB_{yit},
\end{aligned}
$$

where the last step uses that $\ell_{it} = 0$ if $z_{it} < w_t$ by construction in equation (A.10).

Rearranging we obtain the dynamic budget constraint,

$$
dn_{it} = (r_t n_{it} - c_{it} + \ell_{it}(z_{it} - w_t))dt + \ell_{it}\phi_{it}\sigma_{yit} dB_{yit}.
$$

Arrow securities on B_{xit} risk. We can allow entrepreneur to also trade Arrow securities contingent on the realization of their B_{xit} shocks. Their dynamic budget constraint becomes

$$
dn_{it} = (r_t n_{it} - c_{it} + \ell_{it}(z_{it} - w_t))dt + \ell_{it}\sigma_{yit}\phi_{it} dB_{yit} + \sigma_{ait} dB_{xit}.
$$

The main result is that the same (c_i, ℓ_i) as above are still optimal, and entrepreneurs don't trade the Arrow securities, $\sigma_{ait} = 0$.

The same reasoning as above shows that any plan must satisfy the intertemporal budget constraint

$$
\mathbb{E}\left[\int_0^\infty \xi_{it} c_{it} dt\right] \leq n_{i0},
$$

where the definition of ξ_{it} is unchanged. Since the optimal plan solves the relaxed problem, it is unchanged by the introduction of Arrow securities, and can be implemented with $\sigma_{ait} = 0$. Entrepreneurs do not wish to hedge the B_{xit} risk that only affects their investment opportunities. This is a feature of log preferences that yield myopic optimization, in addition to the assumption that B_{yi} and B_{xi} are independent, so trading Arrow securities doesn't help the entrepreneur share uninsurable risk.

Taxes. Now we consider entrepreneurs' problem when the planner introduces taxes on labor and savings, $\tau_\ell(x_{it})$, $\tau_n(x_{it})$ and $\sigma_n(x_{it})$, rebated to workers. In this case, entrepreneurs are not allowed to trade Arrow securities to undo the effect of $\sigma_{nt}(x_{it})$. The budget constraint is

$$dn_{it} = (n_{it}(r_t - \tau_{nt}(x_{it})) - c_{it} + \ell_{it}(z_{it} - \tilde{w}_t(1 + \tau_{\ell t}(z_{it})))) \, dt + n_{it}\sigma_{nt}(x_{it})dB_{xit} + \ell_{it}\sigma_y \phi dB_{yit}.$$

Define

$$d\xi_{it}/\xi_{it} = -(r_t - \tau_{nt}(x_{it}) - \sigma_{nt}^2(x_{it}))dt - \frac{(z_{it} - \tilde{w}_t(1 + \tau_{\ell t}(z_{it}))^+}{\phi_{it}\sigma_{yit}} dB_{yit} - \sigma_{nt}(x_{it})dB_{xit}.$$

The same reasoning as above shows any plan satisfies the intertemporal budget constraint

$$\mathbb{E}\left[\int_0^\infty \xi_{it} c_{it} dt\right] \leq n_{i0}.$$

The relaxed problem has necessary and sufficient condition

$$e^{-\rho_e t} c_{it}^{-1} = \lambda \xi_{it},$$

and λ set so the intertemporal budget constraint hold with equality. This implies c_i is an Ito process

$$dc_{it}/c_{it} = \mu_{cit} dt + \sigma_{cit} dB_{yit} + \sigma_{cit}^x dB_{xit}.$$

$$\mu_{cit} = r_t - \tau_{nt}(x_{it}) - \rho_e + \sigma_{cit}^2,$$

$$\sigma_{cit} = \frac{(z_{it} - \tilde{w}_t(1 + \tau_{\ell t}(z_{it}))^+}{\phi_{it}\sigma_{yit}},$$

$$\sigma_{cit}^x = \sigma_n(x_{it}).$$

To implement it, we set ℓ_{it} according to (A.10) and $n_{it} = c_{it}/\rho_e$ and compute it's law of motion,

$$
\begin{aligned}
dn_{it}/n_{it} &= \left(r_t - \tau_{nt}(x_{it}) - \rho_e + \sigma_{cit}^2\right) dt + \sigma_{cit} dB_{yit} + \sigma_{nt}(x_{it}) dB_{xit}, \\
&= \left(r_t - \tau_{nt}(x_{it}) - c_{it}/n_{it} + \frac{\ell_{it}}{n_{it}} \phi_{it}\sigma_{yit} \frac{(z_{it} - \tilde{w}_t(1+\tau_{\ell t}(z_{it}))^+}{\phi_{it}\sigma_{yit}}\right) dt + \frac{\ell_{it}}{n_{it}} \phi_{it}\sigma_{yit} dB_{yit} + \sigma_{nt}(x_{it}) dB_{xit}, \\
&= \left(r_t - \tau_{nt}(x_{it}) - c_{it}/n_{it} + \frac{\ell_{it}}{n_{it}} (z_{it} - \tilde{w}_t(1+\tau_{\ell t}(z_{it}))\right) dt + \frac{\ell_{it}}{n_{it}} \phi_{it}\sigma_{yit} dB_{yit} + \sigma_{nt}(x_{it}) dB_{xit}.
\end{aligned}
$$

Rearranging, we obtain the dynamic budget constraint.

B Worker Optimality

The worker's problem is to choose a non-negative processes (c_w, ℓ) to maximize utility $U_w(c_w, \ell)$ subject to the dynamic budget constraint

$$dn_{wt} = (n_{wt} r_t - c_{wt} + \tilde{w}_t \ell_t + T_t) dt,$$

with initial bonds n_{w0} given and natural borrowing limit $n_{wt} \geq -\int_t^\infty e^{-\int_t^s r_u du} (\tilde{w}_s \ell_s + T_t) ds$. Here $\tilde{w}_t$ is the after-tax wage that the worker receives, and we added the tax rebates T_t that are used by the planner. For the competitive equilibrium without taxes, simply set $T_t = 0$ and $\tilde{w}_t = w_t$.

This is a standard problem. First, for any plan (c_w, ℓ) satisfying the dynamic budget constraint and natural borrowing limit we satisfy the intertemporal budget constraint

$$\int_0^\infty e^{-\int_0^t r_u du} (c_{wt} - \tilde{w}_t \ell_t - T_t) dt \leq n_{w0}. \tag{B.1}$$

To see this compute

$$d(e^{-\int_0^t r_u du} n_{wt}) = e^{-\int_0^t r_u du} \left(-n_{wt} r_t + n_{wt} r_t - c_{wt} + \tilde{w}_t \ell_t + T_t\right) dt.$$

Integrating we obtain

$$\int_t^T e^{-\int_t^s r_u du} (c_{ws} - \tilde{w}_s \ell_s - T_t) ds + e^{-\int_t^T r_u du} n_{wT} = n_{w0}.$$

Use the natural borrowing limit to obtain

$$\int_t^T e^{-\int_t^s r_u du} c_{ws} ds - \int_t^\infty e^{-\int_t^\infty r_u du} \left(\tilde{w}_s \ell_s + T_t\right) ds \leq n_{w0},$$

and taking the limit $T \to \infty$ we obtain the intertemporal budget constraint. We can therefore consider the relaxed problem of maximizing utility subject to the intertemporal budget constraint.

The first-order conditions of the relaxed problem are

$$e^{-\rho_w t} c_{wt}^{-1} = \lambda e^{-\int_0^t r_u du} \tag{B.2}$$

$$e^{-\rho_w t} \ell_t^{1/\psi} = \lambda e^{-\int_0^t r_u du} \tilde{w}_t \Longleftrightarrow \ell_t^{1/\psi} = c_{wt}^{-1} \tilde{w}_t, \tag{B.3}$$

and λ is set so the intertemporal budget constraint (B.1) holds with equality. These conditions are also sufficient for optimality.

From (B.2), taking time-derivatives, we obtain the Euler equation

$$dc_{wt}/c_{wt} = (r_t - \rho_w)dt. \tag{B.4}$$

A plan (c_w, ℓ) satisfying the first-order conditions (B.4) and (B.3) and the intertemporal budget constraint with equality, achieves the optimum of the relaxed problem.

It remains to show this plan can be implemented with the dynamic budget constraint. Set

$$n_{wt} = \int_t^\infty e^{-\int_t^s r_u du} (c_{ws} - \tilde{w}_s \ell_s - T_t) ds, \tag{B.5}$$

which satisfies the initial condition by construction. Taking derivatives with respect to time, we obtain the dynamic budget constraint

$$dn_{wt} = (n_{wt} r_t - c_{wt} + \tilde{w}_t \ell_t + T_t)dt.$$

Since $c_w \geq 0$, the natural borrowing limit is automatically satisfied.

C Construction of the Competitive Equilibrium

The separation result allows us to compute an equilibrium in two steps. In step one, start with a given distribution $\eta_0(x)$. Then, solve the KFE (21) forward to obtain the path of $\eta_t(x)$. Then use equations (18)–(20) to solve for the path of w_t, Z_t, and $\sigma_c(x)$. At this point we must verify that there exists a positive w_t that solves equation (18) using η_t at all t.

Due to the separation result, step one does not involve aggregate output c_t, aggregate employment ℓ_t, or interest rates r_t. In step two, we determine these by using the representative worker's labor supply condition to solve for aggregate employment and the definition of aggregate TFP to compute aggregate output

$$(\overbrace{\eta_{wt} Z_t \ell_t}^{c_{wt}})^{-1} w_t = \ell_t^{1/\psi}, \tag{C.1}$$

$$c_t = Z_t \ell_t. \tag{C.2}$$

Once we have the path of ℓ and c we can construct the rest of the equilibrium. Worker consumption is $c_{wt} = \eta_{wt} c_t$, and their Euler equation determines the interest rate

$$r_t = \rho_w + \frac{\dot{c}_{wt}}{c_{wt}}. \tag{C.3}$$

Finally, to determine each entrepreneur's allocation, we start with η_{it} consistent with the distribution $\eta_0(x)$ and set $c_{i0} = \eta_{i0} c_0$ and $\sigma_{cit} = \sigma_c(x_{it}; \eta_t)$. Entrepreneurs' Euler equation (11) determine their consumption c_{it}. We set the bonds of each entrepreneur $n_{it} = c_{it}/\rho_e$ and the bonds of the worker to $n_{wt} = -\int_0^1 n_{it} di = \eta_{et} \frac{c_t}{\rho_e}$. Equation (13) determines ℓ_{it}.

We need to show that the prices and allocations $\mathcal{C} = (c_w, \ell, c_i, \ell_i, r, w)$ we constructed constitute a competitive equilibrium. To do so we need to show that allocations and prices are in their permissible domain, that markets clear, and that agents are solving their problem taking prices as given. To do so, we assume that $w_t, c_{it}, c_{wt} > 0$, and $\lim_{T\to\infty} e^{-\int_0^T r_u du} n_{wT} = 0$.

C.1 Market Clearing

First, since $w_t, c_{it}, c_{wt} > 0$, we have that $\ell_{it} = \sigma_c(x_{it}, \eta_t)\frac{c_{it}}{\rho_e \phi_{it} \sigma_{yit}} \geq 0$ and $\ell_t = c_{wt}^{-1} w_t > 0$. Furthermore, by construction $\int_0^1 n_{it} di + n_{wt} = 0$.

To show that markets clear we can use that the KFE is derived from the Euler equations of entrepreneurs and the worker, yielding that

$$\frac{\int_{\{z_{it}=z\}} c_{it} di}{c_{wt}} = \frac{\eta_t(x)}{\eta_{wt}}.$$

Integrating across x and using $\eta_{et} = 1 - \eta_{wt}$ and $c_{wt} = (1 - \eta_{et}) Z_t \ell_t$, we obtain

$$\int_0^1 c_{it} di = \eta_{et} \frac{c_{wt}}{1 - \eta_{et}} = \eta_{et} Z_t \ell_t,$$

and therefore

$$\int_0^1 c_{it} di + c_{wt} = Z_t \ell_t.$$

Equations (18) and (19) then ensure that market clearing (equations 5 and 6) hold.

C.2 Optimality

It remains to show that the allocations (c_i, ℓ_i) and (c_w, ℓ) satisfy the dynamic budget constraints, natural borrowing constraints, and are optimal given prices (r, w).

Since we constructed the entrepreneurs' allocations (c_i, ℓ_i) to satisfy the first-order conditions and $c_{i0} = \rho_e n_{i0}$, they satisfy budget constraints and are optimal (see Appendix A). The allocation (c_w, ℓ) satisfies the first order conditions for the worker. To show optimality, we only need to show that it satisfies the worker's intertemporal budget constraint with equality (see Appendix B). We defined $n_{wt} = -\int_0^1 n_{it} di$. We therefore have that

$$
\begin{aligned}
e^{-\int_t^T r_u du} n_{wT} &= -\int_0^1 e^{-\int_t^T r_u du} n_{it} di, \\
&= \int_0^1 \left(-n_{it} + \int_t^T e^{-\int_t^s r_u du} (c_{is} - (z_{is} - w_s)\ell_{is}) ds - \int_t^T e^{-\int_t^s r_u du} \ell_{is} \phi_{it} \sigma_{yit} dB_{iys} \right) di, \\
&= \int_0^1 -n_{it} di + \int_0^1 \int_t^T e^{-\int_t^s r_u du} (c_{is} - (z_{is} - w_s)\ell_{is}) ds di - \int_0^1 \int_t^T e^{-\int_t^s r_u du} \ell_{is} \phi_{it} \sigma_{yit} dB_{iys} di \\
&= \left(n_{wt} + \int_0^1 \int_t^T e^{-\int_t^s r_u du} (c_{is} - (z_{is} - w_s)\ell_{is}) ds di \right).
\end{aligned}
$$

where the last term is eliminated imposing a law of large numbers.[16] We can switch the order of integration using Tonelli to obtain,

$$e^{-\int_t^T r_u du} n_{wT} = n_{wt} + \int_t^T e^{-\int_t^s r_u du} \left(\int_0^1 c_{is} di - \int_0^1 z_{is} \ell_{is} + w_s \int_0^1 \ell_{is} di \right) ds.$$

Use the resource constraints,

$$e^{-\int_t^T r_u du} n_{wT} = n_{wt} + \int_t^T e^{-\int_t^s r_u du} (-c_{ws} + w_s \ell_s) ds.$$

Use that $e^{-\int_t^T r_u du} n_{wT} = \left(\frac{e^{-\rho w t} c_{wt}^{-1}}{c_{w0}^{-1}} \right)^{-1} e^{-\int_0^T r_u du} n_{wT}$ and take the limit $T \to \infty$. Since $\lim_{T\to\infty} e^{-\int_0^T r_u du} n_{wT} = 0$, then

$$\int_t^\infty e^{-\int_t^s r_u du} (c_{ws} - w_s \ell_s) ds = n_{wt}, \tag{C.4}$$

as desired.

C.3 Conclusion

Thus, initial bonds n_{w0} and n_{i0} and allocations and prices $\mathcal{C} = (c_w, \ell, c_i, \ell_i, r, w)$ are a *Competitive Equilibrium*.

[16]This is implied by the LLN that we imposed when writing the goods resource constraint in equation (5) because $e^{-\int_t^s r_u du}$ as a function of s is deterministic and continuous on $[0, T]$.

D Planner's Problem with Uniform Labor Tax

Starting with η_0 and a path for ℓ from the Planner optimization problem, we can construct the allocation following the same steps we used to construct the competitive equilibrium. Now, however, we do not need to use the worker's first-order condition for labor supply to determine labor supply ℓ_t. Instead, we know ℓ_t and we need to find the labor tax τ_t and the initial bond distribution, n_{w0} and n_{i0}, that implement the allocation. We set the labor tax $\tau_{\ell t}$ according to (30) and the rebate is $T_t = \tau_{\ell t}\tilde{w}_t\ell_t$. Bond holdings are $n_{it} = c_{it}/\rho_e$ and $n_{wt} = -\int_0^1 n_{it}di$. In equation (C.4), use that $w_t\ell_t = \tilde{w}_t(1+\tau_{\ell t})\ell_t = \tilde{w}_t\ell_t + T_t$, to obtain the budget constraint of the worker. We only need to check that $\tilde{w}_t, c_{it}, c_{wt}$ are all positive, and that $\lim_{T\to\infty} e^{-\int_0^T r_u du} n_{wT} = 0$.

E Planner's Problem with Saturated Instruments

E.1 Planner's objective

Entrepreneurs' optimality conditions are[17]

$$dc_{it}/c_{it} = (r_t - \tau_{nt}(x_{it}) - \rho_e + \sigma_{ct}^2(x_{it}))dt + \sigma_{ct}(x_{it})dB_{yit} + \sigma_{nt}(x_{it})dB_{xit}, \tag{E.1}$$

$$z_{it} - w_t(1+\tau_{\ell t}(x_{it})) \leq \sigma_{ct}(x_{it})\phi_{it}\sigma_{yit}, \tag{E.2}$$

with equality if $z_{it} > w_t(1+\tau_{\ell t}(x_{it}))$. As in the competitive equilibrium, $c_{it} = \rho_e n_{it}$, and all entrepreneurs with the same state x_{it} choose the same $\sigma_{ct}(x_{it})$ and therefore the same labor (relative to their net worth or consumption), $\ell_{it}/c_{it} = \sigma_{ct}(x_{it})\frac{1}{\rho_e\phi(x_{it})\sigma_y(x_{it})}$.

The planner can control the consumption distribution $\eta_t(x)$ and the share of aggregate employment by entrepreneurs of type x, $\hat{\ell}(x)$. He can therefore directly control TFP,

$$Z_t = \int z(x)\hat{\ell}_t(x)dx, \tag{E.3}$$

[17] In the Euler equation (E.1) there is no precautionary saving term for B_{xit} risk. The reason is that due to the savings tax $\sigma_{nt}(x_{it})$ savings are not risk-free for entrepreneurs. For workers r_t is the return of a risk-free asset, but for entrepreneurs is the return of a risky asset.

but he faces the following constraint arising from incomplete risk sharing,

$$\sigma_{ct}(x) = \frac{\rho_e \phi(x) \sigma_y(x)}{Z_t \eta_t(x)} \hat{\ell}_t(x). \tag{E.4}$$

When the planner makes entrepreneurs hire more labor, he must expose them to more risk. Within entrepreneurs of current type x, consumption and labor will be dispersed due to their previous history of B_{yit} and B_{xit} shocks. While he cannot eliminate the dispersion due to B_{yit} shocks (i.e., he can't complete markets), it is optimal to eliminate the dispersion arising from B_{xit} shocks. To do this, he sets

$$\eta_{it} = \frac{\eta_t(x_{it})}{p(x_{it})} \times \exp\left(\int_0^t \sigma_{cs}(x_{is}) dB_{yis} - \frac{1}{2} \sigma_{cs}^2(x_{is}) ds \right).$$

An entrepreneurs' consumption share is the average for entrepreneurs of his type x_{it}, times noise due to the previous history of uninsurable B_{yi} risk. Taking time derivatives and matching terms with the law of motion for η_{it} derived from the Euler equations, we obtain the savings taxes that accomplish this,

$$\tau_{nt}(x) = \sigma_{ct}^2(x) - \mathcal{L}_{xt} \frac{\eta_t(x)/p(x)}{\eta_{wt}} \frac{\gamma_{wt}}{\gamma_{et}}, \tag{E.5}$$

$$\sigma_{nt}(x) = \frac{\partial_x \left[\frac{\eta_t(x)/p(x)}{\eta_{wt}} \frac{\gamma_{wt}}{\gamma_{et}} \right]}{\frac{\eta_t(x)/p(x)}{\eta_{wt}} \frac{\gamma_{wt}}{\gamma_{et}}} \sigma_x(x). \tag{E.6}$$

We can therefore rewrite the last term in the planner's objective function (24) as follows. Start with

$$\int_0^\infty \Gamma_{et} \left(\int \left(p(x) \log \frac{\eta_t(x)}{p(x)} - \int \log \eta g_t(\eta, x) d\eta \right) dx \right) dt.$$

Use $\log \eta_{it} = \log \frac{\eta_t(x_{it})}{p(x_{it})} + \int_0^t \sigma_{cs} dB_{yis} - \frac{1}{2}\sigma_{cs}^2 ds$. The first term, $p(x) \log \frac{\eta_t(x)}{p(x)}$, cancels out. We are left with the expectation of $\int_0^t \sigma_{cs} dB_{yis} - \frac{1}{2}\sigma_{cs}^2 ds$. Use a law of large numbers to eliminate the stochastic integral. Then switch the order of integration:

$$\int_0^\infty \Gamma_{et} \int_0^t \frac{1}{2} \sigma_{cis}^2 ds dt = \int_0^\infty \Gamma_{et} \frac{1}{2} \frac{1}{\rho_e} \sigma_{cit}^2 dt.$$

The advantage of this expression is that σ_{cit} is the same for all agents with the same x_{it}. So integrating conditional on x first, we obtain the expression for the last term,

$$\int_0^\infty \Gamma_{et}\frac{1}{\rho_e}\frac{1}{2}\left(\int p(x)\sigma_{ct}^2(x)dx\right)dt.$$

Replacing this in the planner's objective we obtain,

$$\overbrace{\int_0^\infty (\Gamma_{wt}+\Gamma_{et})\log\ell_t - \Gamma_{wt}\frac{\ell_t^{1+1/\psi}}{1+1/\psi}dt}^{\text{labor}} + \overbrace{\int_0^\infty (\Gamma_{wt}+\Gamma_{et})\log Z_t dt}^{\text{TFP}} \tag{E.7}$$
$$+\underbrace{\int_0^\infty \Gamma_{wt}\log\eta_{wt} + \Gamma_{et}\left(\int p(x)\log\frac{\eta_t(x)}{p(x)}dx\right)dt}_{\text{workers vs. entrepreneurs}} - \underbrace{\int_0^\infty \Gamma_{et}\frac{1}{\rho_e}\frac{1}{2}\left(\int p(x)\sigma_{ct}^2(x)dx\right)dt}_{\text{entrepreneur risk}}.$$

E.2 Optimal allocation

Aggregate employment ℓ_t appears separately in the planner's objective function and does not enter constraints. The planner therefore chooses the first-best level of employment, $\ell_t = \ell_t^{FB}$.

The FOC for $\eta_t(x)$ is

$$\gamma_{et}\frac{p(x)}{\eta_t(x)} - \gamma_{wt}\frac{1}{1-\eta_{et}} + \gamma_{et}\frac{1}{\rho_e}\sigma_{ct}^2(x)\frac{p(x)}{\eta_t(x)} = 0.$$

After some algebra we obtain

$$\frac{\eta_t(x)}{p(x)} = \frac{\gamma_{et}\left(1+\frac{1}{\rho_e}\sigma_{ct}^2(x)\right)}{\gamma_{wt}+\gamma_{et}\left(1+\frac{1}{\rho_e}\sigma_{ct}^2\right)}, \tag{E.8}$$

where

$$\sigma_{ct}^2 = \int p(x)\sigma_{ct}^2(x)dx$$

is the average uninsurable idiosyncratic risk.

The FOC $\hat{\ell}_t(x)$ is

$$\frac{z(x)}{Z_t} - \gamma_{et}\frac{1}{\rho_e}\sigma_{ct}^2(x)\frac{p(x)}{\hat{\ell}_t(x)} + \gamma_{et}\frac{1}{\rho_e}\frac{z(x)}{Z_t}\sigma_{ct}^2 \leq \lambda_t,$$

with equality if $\hat{\ell}_t(x) > 0$. The first term captures the effect of higher $\hat{\ell}(x)$ on aggregate TFP, while the second term captures the cost of exposing entrepreneurs of type x to more risk. The third term captures the effect of aggregate TFP on risk sharing. For given aggregate employment ℓ_t, higher TFP raises aggregate output and therefore reduces all entrepreneurs' exposure to risk. Multiply throughout by $\hat{\ell}(x)$ so it's an equality for all x, and integrate across x, to obtain $\lambda_t = 1$. We can solve for $\hat{\ell}_t(x)$,

$$\hat{\ell}_t(x) = \left[z(x)\left(\gamma_{wt} + \gamma_{et}\left(1 + \frac{1}{\rho_e}\sigma_{ct}^2\right)\right) - Z_t\right]^+ \frac{Z_t \eta_t(x)}{\rho_e\left(\phi(x)\sigma_y(x)\right)^2 \gamma_{et}\frac{p(x)}{\eta_t(x)}}.$$

E.3 Implementation

Tax on labor. The derivation of the average labor tax is in the main body of the paper. Here we study how the labor tax depends on the entrepreneurs' type. In the competitive equilibrium the equilibrium condition for entrepreneurs who produce ($\ell(x) > 0$) is

$$z(x) - \sigma_{ct}(x)\sigma_y(x)\phi(x) = w_t(1 + \tau_{\ell t}(x)).$$

The condition in the planner's allocation is

$$z(x)\left(\gamma_{wt} + \gamma\Gamma_{et}\left(1 + \frac{1}{\rho_e}\sigma_{ct}^2\right)\right) - \sigma_{ct}(x)\phi(x)\sigma_y(x)\gamma_{et}\frac{p(x)}{\eta_t(x)} = Z_t$$

We can then back out the cross-section of the optimal tax by subtracting one from the other and rearranging terms,

$$w_t(1 + \tau_{\ell t}(x)) - Z_t = -\gamma_{et}\frac{1}{\rho_e}\sigma_{ct}^2 z(x) + \sigma_{ct}(x)\phi(x)\sigma_y(x)\left[\gamma_{et}\frac{p(x)}{\eta_t(x)} - 1\right]. \tag{E.9}$$

Tax on savings. The tax on savings is already derived in (E.5) and (E.6).

F Case with $\rho_e = \rho_w$

For part of Section 39 we assumed $\rho_e > \rho_w$. This is done to obtain a non-degenerate stationary wealth distribution in the competitive equilibrium, avoiding entrepreneurs' precautionary saving leading them to accumulate all the wealth. But our results are robust to the case with $\rho_e = \rho_w$. In this case, entrepreneurs' precautionary saving motive σ_{cit}^2—which is always bounded away from zero for the positive measure of entrepreneurs who produce—makes their expected consumption growth larger than workers' consumption growth (compare Euler equations (9) and (11)).[18] This means when $\rho_e = \rho_w$, in the long run entrepreneurs accumulate all the wealth, i.e., $\eta_{et} \to 1$.

The expressions for finite t in the main body of the paper are still valid. The only results that are affected are those pertaining to the long run. Although their precise statement must be adjusted, the spirit remains the same.

Workers' Consumption Share is Less than their Pareto Weight in the Long Run. In the baseline case with $\rho_e > \rho_w$, entrepreneurs' Pareto weight $\gamma_{et} \to 0$ while their share of consumption does not because of their precautionary saving. When $\rho_e = \rho_w$, entrepreneurs' Pareto weights are constant and less than 1, $\gamma_{et} = \gamma_{e0} \in (0, 1)$. Just as in the $\rho_e > \rho_w$ case, entrepreneurs precautionary saving makes $\eta_{et} \to 1$, larger than their Pareto weight in the long run.

In the Long Run the Optimal Labor Tax is Positive (Not a Subsidy). In Section 3.3 we derive equation (39) for the optimal tax in the long run, under the assumption that η_t converges to a non-degenerate stationary distribution. A key result is that, in the long run, the planner uses a tax on labor,

$$1 + \tau_\ell^* = \frac{w^* \ell^*}{c_w^*} > 1,$$

in contrast to the typical optimal subsidy in the monopolistic competition markup literature. This result is still true if $\rho_e = \rho_w$, and in a more extreme form. First, the general expression for the optimal tax (not just in the long run) is

$$1 + \tau_t = \frac{1}{1 + \pi_t} \frac{\gamma_{wt}}{\eta_{wt}}.$$

[18]If $w = \bar{z}$ then the left hand side of equation (18) would be zero, which means w must be bounded away below $\bar{z}$. Equation (20) shows σ_{cit} is thus bounded away from zero for all $z_{it} > w_t$.

Use $1 + \pi_t = Z_t / w_t$ and $\eta_{wt} = c_{wt}/(Z_t \ell_t)$ to write,

$$1 + \tau_{\ell t} = \frac{w_t \ell_t}{c_{wt}} \gamma_{wt}.$$

When $\rho_e > \rho_w$, the term $\gamma_{wt} \to 1$. If $(w_t \ell_t)/c_{wt}$ converges, it must be to something larger than one. In the long run, workers have positive debt so their consumption must be less than their labor income (the tax on labor and the lump-sum transfers cancel out). From this we get get the result that $\tau_t \to \tau^* > 0$.

When $\rho_e = \rho_w$, the term $\gamma_{wt} = \gamma_{w0} \in (0, 1)$ for all t. The planner always sets labor to the first-best level and aggregate TFP converges to $Z(\eta^*)$, so output converges to $c^* > 0$. Since $\eta_{et} \to 1$, then $n_{et} \to n_{e^*} = \frac{c^*}{\rho_e} > 0$ and $c_{wt} \to 0$. By bond market clearing, $n_w^* = -n_e^* < 0$, so workers natural borrowing limit must be strictly positive. If $w_t \to 0$, then the natural borrowing constraint would be zero (because ℓ^{FB} is finite), which cannot be. Therefore $(w_t \ell_t)/c_{wt} \to \infty$. This yields $\tau_{\ell t} \to \infty$. In the long run, the optimal labor tax is not only positive (a tax), it becomes unboundedly large. As workers become immiserated, the income effect raises labor supply, and an extreme tax is necessary to implement the finite first-best employment.

G Derivation of KFE for $\eta_t(z)$

We explicitly derive the KFE that characterizes the distribution of consumption shares under the assumptions we make in Appendix H, so that productivity follows the process

$$dz_{it} = \mu_z(z_{it})\, dt + \sigma_z(z_{it})\, dB_{zit}$$

where B_{zit} is a Brownian motion.

G.1 Derive process for η_{it}

We need to keep track of (z_i, η_i) where $\eta_{it} = c_{it}/c_t$. We need to compute the growth rate of aggregate consumption c_t:

$$c_{wt}\mu_{cwt} + \int c_{it}\mu_{cit} di = \mu_{ct} c_t$$

$$\eta_{wt}\mu_{cwt} + \int \eta_t(z)\mu_{ct}(z)dz = \mu_{ct}$$

(using the fact that growth rates depend only on z type).

Now use the Euler equations to rewrite this as

$$\eta_{wt}(r_t - \rho_w) + \int \eta_t(z)\left(r_t - \rho_e + \sigma_{ct}(z)^2\right)dz = \mu_{ct}$$

$$\left(\underbrace{\eta_{wt} + \int \eta_t(z)dz}_{=1}\right) r_t - \underbrace{\left(\eta_{wt}\rho_w + \rho_e \int \eta_t(z)dz\right)}_{=\bar{\rho}_t} + \int \eta_t(z)\sigma_{ct}(z)^2 dz = \mu_{ct}$$

$$\mu_{ct} = r_t - \bar{\rho}_t + \int \eta_t(z)\sigma_{ct}(z)^2 dz.$$

So now we can compute the law of motion of $\eta_{it} = c_{it}/c_t$:

$$d(c_{it}/c_t) = \left(\frac{c_{it}}{c_t}\right)(\mu_{cit} - \mu_{ct})\,dt + \left(\frac{c_{it}}{c_t}\right)\sigma_{cit}dB_{yit}$$

$$d(c_{it}/c_t) = \left(\frac{c_{it}}{c_t}\right)\underbrace{\left(\bar{\rho}_t - \rho_e + \sigma_{ct}(z)^2 - \int \eta_t(z)\sigma_{ct}(z)^2 dz\right)}_{\mu_\eta(z,t)} dt$$

$$+ \left(\frac{c_{it}}{c_t}\right)\sigma_{ct}(z)dB_{yit},$$

or, more compactly,

$$d\eta_{it} = \eta_{it}\mu_\eta(z,t)dt + \eta_{it}\sigma_{ct}(z)dB_{yit}.$$

G.2 Derive KFE

Let $D(\eta, z, t)$ be the joint distribution of η, z. $D(\eta, z, t)$ is the measure of entrepreneurs with $z_i = z$ and $\eta_i = \eta$. This distribution has a convenient mapping into $\eta(z,t)$ as

$$\eta(z,t) = \int D(\eta, z, t)\eta d\eta$$

Let $X = (z_i, \eta_i)$

$$dX_t = \begin{bmatrix} \mu_z(z) \\ \mu_\eta(z,t)\eta \end{bmatrix} dt + \begin{bmatrix} \sigma_z(z) & 0 \\ 0 & \sigma_c(z,\eta_t)\eta \end{bmatrix} \begin{bmatrix} dB_{zt} \\ dB_{yt} \end{bmatrix}.$$

To simplify the notation, we can write more generally

$$dX_t = \begin{bmatrix} \mu_z(z) \\ \mu_\eta(z,t)\eta \end{bmatrix} dt + \begin{bmatrix} \sigma_{z1}(z) & \sigma_{z2}(z) \\ \sigma_{c1}(z,t)\eta & \sigma_{c2}(z,t)\eta \end{bmatrix} \begin{bmatrix} dB_{zt} \\ dB_{yt} \end{bmatrix}.$$

We derive the KFE for $D(\eta, z, t)$, starting from the first order terms[19]

$$\begin{aligned} \partial D(\eta,z,t) &= -\frac{\partial}{\partial z}\left[\mu_z(z)D(\eta,z,t)\right] - \frac{\partial}{\partial \eta}\left[\mu_\eta(z,t)\eta D(\eta,z,t)\right] + \text{ 2nd order terms} \\ &= -\left[\mu_z'(z)D(\eta,z,t) + \mu_z(z)D_z'(\eta,z,t)\right] \\ &\quad -\left[\mu_\eta(z)D(\eta,z,t) + \mu_\eta(z)\eta D_\eta'(\eta,z,t)\right] + \text{ 2nd order terms.} \end{aligned}$$

Now, we define the second order terms. The variance covariance matrix is

$$\Lambda = \begin{bmatrix} \sigma_{z1}(z) & \sigma_{z2}(z) \\ \sigma_{c1}(z,t)\eta & \sigma_{c2}(z,t)\eta \end{bmatrix} \begin{bmatrix} \sigma_{z1}(z) & \sigma_{z2}(z) \\ \sigma_{c1}(z,t)\eta & \sigma_{c2}(z,t)\eta \end{bmatrix}'$$

$$\Lambda = \begin{bmatrix} \sigma_{z1}^2(z) + \sigma_{z2}^2(z) & \eta\sigma_{z1}(z)\sigma_{c1}(z) + \eta\sigma_{z2}(z)\sigma_{c2}(z) \\ \eta\sigma_{z1}(z)\sigma_{c1}(z) + \eta\sigma_{z2}(z)\sigma_{c2}(z) & \sigma_{c1}^2(z)\eta^2 + \sigma_{c2}^2(z)\eta^2 \end{bmatrix},$$

$$\text{2nd order terms} = \frac{1}{2}\frac{\partial^2}{\partial z^2}[\Lambda_{11}D(\eta,z,t)] + \frac{1}{2}\frac{\partial^2}{\partial \eta^2}[\Lambda_{22}D(\eta,z,t)] + \frac{\partial^2}{\partial \eta z}[\Lambda_{12}D(\eta,z,t)].$$

[19]Since $d\eta_{it}/\eta_{it} = dc_{it}/c_{it}$, we have $d\eta_i = \mu_c\eta_i dt + \sigma_c\eta_i dB_t$.

Further evaluating the matrix by elements, the first term is:

$$\begin{aligned}
\frac{1}{2}\frac{\partial^2}{\partial z^2}[\Lambda_{11}D(\eta,z,t)] &= \frac{1}{2}\frac{\partial^2}{\partial z^2}[(\sigma_{z1}^2(z)+\sigma_{z2}^2(z))D(\eta,z,t)] \\
&= \frac{1}{2}\frac{\partial}{\partial z}[2(\sigma_{z1}(z)\sigma_{z1}'(z)+\sigma_{z2}(z)\sigma_{z2}'(z))D(\eta,z,t)+(\sigma_{z1}^2(z)+\sigma_{z2}^2(z))D_z(\eta,z,t)] \\
&= \frac{\partial}{\partial z}[(\sigma_{z1}(z)\sigma_{z1}'(z)+\sigma_{z2}(z)\sigma_{z2}'(z))D(\eta,z,t)] \\
&\quad + \frac{1}{2}\frac{\partial}{\partial z}[(\sigma_{z1}^2(z)+\sigma_{z2}^2(z))D_z(\eta,z,t)] \\
&= [\left(\sum_j \sigma_{zj}'(z)^2+\sigma_{zj}(z)\sigma_{zj}''(z)\right)D(\eta,z,t)]+[\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)D_z(\eta,z,t)] \\
&\quad + [\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)D_z(\eta,z,t)]+\frac{1}{2}[(\sigma_{z1}^2(z)+\sigma_{z2}^2(z))D_{zz}(\eta,z,t)] \\
&= [\left(\sum_j \sigma_{zj}'(z)^2+\sigma_{zj}(z)\sigma_{zj}''(z)\right)D(\eta,z,t)]+2[\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)D_z(\eta,z,t)] \\
&\quad + \frac{1}{2}[(\sigma_{z1}^2(z)+\sigma_{z2}^2(z))D_{zz}(\eta,z,t)].
\end{aligned}$$

The second term is:

$$\begin{aligned}
\frac{1}{2}\frac{\partial^2}{\partial \eta^2}[\Lambda_{22}D(\eta,z,t)] =& \frac{1}{2}\frac{\partial^2}{\partial \eta^2}[(\sigma_{c1}^2(z)\eta^2+\sigma_{c2}^2(z)\eta^2)D(\eta,z,t)] \\
=&(\sigma_{c1}^2(z)+\sigma_{c2}^2(z))\frac{1}{2}\frac{\partial^2}{\partial \eta^2}[\eta^2 D(\eta,z,t)] \\
=&(\sigma_{c1}^2(z)+\sigma_{c2}^2(z))\frac{1}{2}\frac{\partial}{\partial \eta}[2\eta D(\eta,z,t)+\eta^2 D_\eta(\eta,z,t)] \\
=&(\sigma_{c1}^2(z)+\sigma_{c2}^2(z))\left[\frac{\partial}{\partial \eta}[\eta D(\eta,z,t)]+\frac{1}{2}\frac{\partial}{\partial \eta}[\eta^2 D_\eta(\eta,z,t)]\right] \\
=&(\sigma_{c1}^2(z)+\sigma_{c2}^2(z))\left[[D(\eta,z,t)+\eta D_\eta(\eta,z,t)]+\eta D_\eta(\eta,z,t)+\frac{1}{2}[\eta^2 D_{\eta\eta}(\eta,z,t)]\right] \\
=&(\sigma_{c1}^2(z)+\sigma_{c2}^2(z))\left[[D(\eta,z,t)+2\eta D_\eta(\eta,z,t)]+\frac{1}{2}[\eta^2 D_{\eta\eta}(\eta,z,t)]\right].
\end{aligned}$$

The third term is:

$$
\begin{aligned}
\frac{\partial^2}{\partial\eta\partial z}[\Lambda_{12}D(\eta,z,t)] &= \frac{\partial^2}{\partial\eta\partial z}[[\eta\sigma_{z1}(z)\sigma_{c1}(z)+\eta\sigma_{z2}(z)\sigma_{c2}(z)]D(\eta,z,t)] \\
&= \frac{\partial^2}{\partial\eta\partial z}[\sum_{j=1}^{2}(\sigma_{zj}(z)\sigma_{cj}(z))\eta D(\eta,z,t)] \\
&= \frac{\partial}{\partial z}[\sum_{j=1}^{2}(\sigma_{zj}(z)\sigma_{cj}(z))D(\eta,z,t)] + \frac{\partial}{\partial z}[\sum_{j=1}^{2}(\sigma_{zj}(z)\sigma_{cj}(z))\eta D_\eta(\eta,z,t)] \\
&= \sum_{j=1}^{2}\left[\sigma'_{zj}(z)\sigma_{cj}(z)+\sigma_{zj}(z)\sigma'_{cj}(z)\right]D(\eta,z,t) + \sum_{j=1}^{2}[\sigma_{zj}(z)\sigma_{cj}(z)D_z(\eta,z,t)] \\
&+ \sum_{j=1}^{2}\left[\sigma'_{zj}(z)\sigma_{cj}(z)+\sigma_{zj}(z)\sigma'_{cj}(z)\right]\eta D_\eta(\eta,z,t) \\
&+ \sum_{j=1}^{2}[\sigma_{zj}(z)\sigma_{cj}(z))\eta D_{\eta z}(\eta,z,t)]\,.
\end{aligned}
$$

Combining all first and second order terms yields

$$
\begin{aligned}
\partial D(\eta,z,t) = &-\left[\mu'_z(z)D(\eta,z,t)+\mu_z(z)D'_z(\eta,z,t)\right]-\left[\mu_\eta(z)D(\eta,z,t)+\mu_\eta(z)\eta D'_\eta(\eta,z,t)\right] \\
&+\left(\sum_j \sigma'_{zj}(z)^2+\sigma_{zj}(z)\sigma''_{zj}(z)\right)D(\eta,z,t)+2\left(\sum_j\sigma_{zj}(z)\sigma'_{zj}(z)\right)D_z(\eta,z,t) \\
&+\frac{1}{2}[(\sigma_{z1}^2(z)+\sigma_{z2}^2(z))D_{zz}(\eta,z,t)] \\
&+(\sigma_{c1}^2(z)+\sigma_{c2}^2(z))\left[[D(\eta,z,t)+2\eta D_\eta(\eta,z,t)]+\frac{1}{2}[\eta^2 D_{\eta\eta}(\eta,z,t)]\right] \\
&+\sum_{j=1}^{2}\left[\sigma'_{zj}(z)\sigma_{cj}(z)+\sigma_{zj}(z)\sigma'_{cj}(z)\right]D(\eta,z,t)+\sum_{j=1}^{2}[\sigma_{zj}(z)\sigma_{cj}(z)D_z(\eta,z,t)] \\
&+\sum_{j=1}^{2}\left[\sigma'_{zj}(z)\sigma_{cj}(z)+\sigma_{zj}(z)\sigma'_{cj}(z)\right]\eta D_\eta(\eta,z,t)+\sum_{j=1}^{2}[\sigma_{zj}(z)\sigma_{cj}(z))\eta D_{\eta z}(\eta,z,t)]\,.
\end{aligned}
$$

To obtain $\eta(z,t)$, use that

$$\begin{aligned}
\eta(z,t) &= \int \eta D(\eta,z,t)d\eta \\
\partial_t \eta(z,t) &= \int \eta \partial D(\eta,z,t)d\eta \\
\eta_z(z,t) &= \int \eta D_z(\eta,z,t)d\eta \\
\eta_{zz}(z,t) &= \int \eta D_{zz}(\eta,z,t)d\eta.
\end{aligned}$$

Then,

$$\begin{aligned}
\partial\eta(z,t) = &- \left[\mu_z'(z)\eta(z,t) + \mu_z(z)\eta_z'(\eta,z,t)\right] - \left[\mu_\eta(z)\eta(z,t) + \mu_\eta(z)\int \eta^2 D_\eta'(\eta,z,t)d\eta\right] \\
&+ \left(\sum_j \sigma_{zj}'(z)^2 + \sigma_{zj}(z)\sigma_{zj}''(z)\right)\eta(z,t) + 2\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)\eta_z'(z,t) \\
&+ \frac{1}{2}[(\sigma_{z1}^2(z) + \sigma_{z2}^2(z))\eta_{zz}''(\eta,z,t)] \\
&+ (\sigma_{c1}^2(z) + \sigma_{c2}^2(z))\left[[\eta(z,t) + 2\int \eta^2 D_\eta(\eta,z,t)d\eta] + \frac{1}{2}[\int \eta^3 D_{\eta\eta}(\eta,z,t)d\eta]\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right]\eta(z,t) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z)\eta_z'(z,t)\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right]\int \eta^2 D_\eta(\eta,z,t)d\eta + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z))\int \eta^2 D_{\eta z}(\eta,z,t)d\eta\right].
\end{aligned}$$

Conjecture that $\eta^2 D(z,\eta,t)\big|_0^1 = \lim_{\eta\to 1}\eta^2 D(z,\eta,t) - \lim_{\eta\to 0}\eta^2 D(z,\eta,t) = 0 - 0 = 0$. Then, using integration by parts,

$$\int_0^1 \eta^2 D_\eta(z,\eta,t)d\eta = \eta^2 D(z,\eta,t)\big|_0^1 - 2\int_0^1 \eta D(z,\eta,t)d\eta = -2\eta(z,t).$$

Substitute this into the KFE to obtain

$$
\begin{aligned}
\partial\eta(z,t) =& - \left[\mu_z'(z)\eta(z,t) + \mu_z(z)\eta_z'(\eta,z,t)\right] - \left[\mu_\eta(z)\eta(z,t) + \mu_\eta(z)(-2\eta(z,t))\right] \\
&+ \left(\sum_j \sigma_{zj}'(z)^2 + \sigma_{zj}(z)\sigma_{zj}''(z)\right)\eta(z,t) + 2\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)\eta_z'(z,t) \\
&+ \frac{1}{2}[(\sigma_{z1}^2(z) + \sigma_{z2}^2(z))\eta_{zz}''(\eta,z,t)] \\
&+ (\sigma_{c1}^2(z) + \sigma_{c2}^2(z))\left[[\eta(z,t) + 2(-2\eta(z,t))] + \frac{1}{2}[\int \eta^3 D_{\eta\eta}(\eta,z,t)d\eta]\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right]\eta(z,t) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z)\eta_z'(z,t)\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right](-2\eta(z,t)) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z))\int \eta^2 D_{\eta z}(\eta,z,t)d\eta\right].
\end{aligned}
$$

Conjecture that $\eta^3 D_\eta(z,\eta,t)\big|_0^1 = \lim_{\eta\to 1}\eta^3 D_\eta(z,\eta,t) - \lim_{\eta\to 0}\eta^3 D_\eta(z,\eta,t) = 0 - 0 = 0$. Then, using integration by parts,

$$
\int_0^1 \eta^3 D_{\eta\eta}(\eta,z,t)d\eta = \eta^3 D_\eta(z,\eta,t)\big|_0^1 - 3\int_0^1 \eta^2 D_\eta(\eta,z,t)d\eta = 6\eta(z,t).
$$

Again, substitute this into the KFE to obtain

$$
\begin{aligned}
\partial\eta(z,t) =& - \left[\mu_z'(z)\eta(z,t) + \mu_z(z)\eta_z'(\eta,z,t)\right] - \left[\mu_\eta(z)\eta(z,t) + \mu_\eta(z)(-2\eta(z,t))\right] \\
&+ \left(\sum_j \sigma_{zj}'(z)^2 + \sigma_{zj}(z)\sigma_{zj}''(z)\right)\eta(z,t) + 2\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)\eta_z'(z,t) \\
&+ \frac{1}{2}[(\sigma_{z1}^2(z) + \sigma_{z2}^2(z))\eta_{zz}''(\eta,z,t)] \\
&+ (\sigma_{c1}^2(z) + \sigma_{c2}^2(z))\left[[\eta(z,t) + 2(-2\eta(z,t))] + \frac{1}{2}[6\eta(z,t)]\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right]\eta(z,t) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z)\eta_z'(z,t)\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right](-2\eta(z,t)) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z)\int \eta^2 D_{\eta z}(\eta,z,t)d\eta\right].
\end{aligned}
$$

A similar argument can be applied to obtain

$$\int_0^1 \eta^2 D_{\eta z}(\eta, z, t) d\eta = -2\eta_z'(z, t),$$

which implies that

$$\begin{aligned}
\partial\eta(z,t) =& -\left[\mu_z'(z)\eta(z,t) + \mu_z(z)\eta_z'(\eta,z,t)\right] - \left[\mu_\eta(z)\eta(z,t) + \mu_\eta(z)(-2\eta(z,t))\right] \\
&+ \left(\sum_j \sigma_{zj}'(z)^2 + \sigma_{zj}(z)\sigma_{zj}''(z)\right)\eta(z,t) + 2\left(\sum_j \sigma_{zj}(z)\sigma_{zj}'(z)\right)\eta_z'(z,t) \\
&+ \frac{1}{2}[(\sigma_{z1}^2(z) + \sigma_{z2}^2(z))\eta_{zz}''(\eta,z,t)] \\
&+ (\sigma_{c1}^2(z) + \sigma_{c2}^2(z))\left[[\eta(z,t) + 2(-2\eta(z,t))] + \frac{1}{2}[6\eta(z,t)]\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right]\eta(z,t) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z)\eta_z'(z,t)\right] \\
&+ \sum_{j=1}^{2}\left[\sigma_{zj}'(z)\sigma_{cj}(z) + \sigma_{zj}(z)\sigma_{cj}'(z)\right](-2\eta(z,t)) + \sum_{j=1}^{2}\left[\sigma_{zj}(z)\sigma_{cj}(z))(-2\eta_z'(z,t))\right].
\end{aligned}$$

Grouping terms and using $\sigma_{c1}(z) = \sigma_{z2}(z) = 0$ yields

$$\begin{aligned}
\partial\eta(z,t) =& \eta(z,t)\left[-\mu_z'(z) + \mu_\eta(z) + \left(\sigma_z'(z)^2 + \sigma_z(z)\sigma_z''(z)\right)\right] \\
&+ \eta_z'(z,t)\left[-\mu_z(z) + 2\sigma_z(z)\sigma_z'(z)\right] \\
&+ \eta_{zz}''(\eta,z,t)\left[\frac{1}{2}\sigma_z^2(z)\right].
\end{aligned}$$

where $\mu_\eta(z,t)$ is determined in equilibrium and is given by:

$$\mu_\eta(z,t) = \bar{\rho}_t - \rho_e + \sigma_{ct}(z)^2 - \int \eta_{zt}\sigma_{ct}(z)^2 dz.$$

H Numerical Example

H.1 Calibration

Parameter	Description	Value	Moment	Value
ρ	Workers' discount factor	0.035	Interest rate	3.5%
ρ_e	Entrepreneurs' discount factor	0.170	Wealth/Output	1.5
θ	Mean reversion of z_{it}	0.30	Autocorrelation of z_{it}	0.92
$\sigma_{\log z}$	Std. deviation of B_{zt}	0.04	Std. dev. of z_{it} shocks	0.04
σ_y	Std. deviation of B_{yt}	0.47	Average markup	0.25
ψ	Frisch Elasticity	1		
ϕ	Exposure to idiosyncratic risk	1		

Table H.1: Targeted Moments and Parameter Values for the Numerical Illustration.

In order to solve the model numerically, we impose more structure on the stochastic process for x_{it}, where realizations of x_{it} are realizations of firms' productivity. The law of motions of entrepreneurs' net worth is characterized by

$$dn_{it} = (r_t n_{it} - c_{it} + (z_{it} - w_t)\,\ell_{it})\,dt + (\phi\sigma_y)\,\ell_{it} z_{it} dB_{yit}$$

$$d\log z_{it} = \theta\,(\log \bar{z} - \log z_{it})\,dt + \sigma_{\log z} dB_{zt}$$

where B_{yit} and B_{zit} are two Brownian motions specific to the entrepreneur. We assume that $\log z$ follows a standard Ornstein-Uhlenbeck process. Moreover, we normalize $\log(\bar{z}) = -\frac{1}{\theta}\frac{1}{2}\sigma^2_{\log z}$, so that the implied stochastic process for z takes a simple form:

$$dz_{it} = -\theta z_{it} \log(z_{it})\,dt + z_{it}\sigma_{\log z} dB_{zt}$$

Therefore, the stochastic process driving fluctuations in firms' output and productivity is disciplined by a set of parameters $(\sigma_y, \sigma_{\log z}, \theta)$.

Parameter values are reported in Table H.1. We interpret a unit of time in our model as corresponding to one year. We set the workers discount factor ρ equal to 0.035, to match an annual interest rate of 3.5 percent. We calibrate the entrepreneurs discount factor ρ^e in order to match an average business wealth to output ratio equal to 1.5, as in Boar, Gorea, and Midrigan (2025).

These parameter values imply that in steady state workers consume approximately 93% of labor income, so that $c_w = 0.93w\ell$. We set the parameters θ, σ_{logz} to match the same autocorrelation and standard deviation of firms' productivity as in Boar, Gorea, and Midrigan (2025). Since the moments reported in Boar, Gorea, and Midrigan (2025) are for a sample of private firms, we set $\phi = 1$.

We set the Frisch elasticity ψ equal to one, a conventional value in the literature that lies between micro and macro estimates Chetty, Guren, Manoli, and Weber (2011). Finally, we set the level of idiosyncratic risk to match an average markup of 25%, in line with estimates from Edmond, Midrigan, and Xu (2023)

H.2 Numerical algorithm

Competitive Equilibrium. We use the separation result to compute the competitive equilibrium in two steps. First, we use the KFE and the good market-clearing condition to solve for the distribution of consumption shares $\eta_t(z)$ and for the wage w_t. Then, it is easy to compute aggregate labor and consumption using the labor market-clearing condition and the first-order condition for labor supply.

In order to solve for the distribution of consumption shares, we construct a grid for productivity $\mathbf{z}$ that takes values between $\underline{z}$ and $\bar{z}$. We denote by η the vector obtained by evaluating $\eta(z)$ at $\mathbf{z}$.

Let $\mathbf{D_f}$ be the forward operator and $\mathbf{D_b}$ be the backward operator, defined as:

$$\mathbf{D_f} = \begin{pmatrix} -1/dz & 1/dz & 0 & \dots & 0 \\ 0 & -1/dz & 1/dz & 0 & \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \dots & & -1/dz & 1/dz \\ 0 & \dots & & 0 & -1 \end{pmatrix}$$

$$\mathbf{D_b} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -1/dz & 1/dz & 0 & & \\ \vdots & -1/dz & 1/dz & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \dots & & -1/dz & 1/dz \end{pmatrix}$$

For first derivatives, we use the operator $\mathbf{D^1}$ defined as:

$$\mathbf{D^1} = \mathbf{1}(z > \bar{z})\mathbf{D_b} + \mathbf{1}(z < \bar{z})\mathbf{D_f}.$$

For second derivatives, we use the operator $\mathbf{D^2}$ defined as:

$$\mathbf{D^2} = \mathbf{D_f}\mathbf{D_b}.$$

To handle the finite domain, we follow Moll (2014) to ensure that the support of the distribution is bounded. We assume two reflecting barriers at the bounds $\underline{z}$ and $\bar{z}$.

We initialize the algorithm with a value for the initial distribution of consumption shares $\eta_{\mathbf{0}}$. Then, at each step of the algorithm:

1. We use the good market clearing condition to find the wage

2. We update the distribution $\eta_{\mathbf{1}}$ using the Runge-Kutta method

3. In order to minimize numerical errors when we iterate forward, we use our equation for $d\eta_e$

to normalize the distribution of consumption shares at each step, such that:

$$\eta_{\mathbf{1}}{}^{norm} = \eta_{\mathbf{1}} \times \frac{\sum \eta_{\mathbf{0}} + d\eta_e}{\sum \eta_{\mathbf{1}}}$$

4. We update the distribution by setting $\eta_{\mathbf{0}} = \eta_{\mathbf{1}}{}^{norm}$ and iterate to convergence

We simulate the algorithm until the economy converges to a steady state, that is when $\eta_{\mathbf{1}} \approx \eta_{\mathbf{0}}$.

Once we have solved for the entire path of the distribution of consumption shares and wages, we can easily compute aggregate labor, consumption, and productivity as

$$\begin{aligned} Z_t &= \left(\tfrac{1}{\rho^e} \int_{w_t}^{\bar{z}} \tfrac{z - w_t}{\sigma_y(z)} \tfrac{\eta_{zt}(z)}{\sigma_y(z)} dz \right)^{-1} && \text{(Aggregate TFP)} \\ \ell_t^{1/\psi} &= \left(\eta_{wt} Z_t \ell_t\right)^{-1} w_t && \text{(Aggregate labor)} \\ c_t &= \ell_t Z_t && \text{(Aggregate consumption)} \end{aligned}$$

Planner's Problem. The numerical algorithm that we use to solve the planner's problem is very similar to the one for the competitive equilibrium.

Once we have solved the entire path of the distribution of consumption shares and wages, as in the competitive equilibrium, we can easily find the optimal labor tax as a function of consumption shares, markups, and Pareto weights. Then, given the optimal labor tax and consumption shares, we solve for aggregate labor and consumption.

In order to assess welfare gains, we need to evaluate the objective function of the social planner. We use the equation for the transformed planner's problem:

$$\overbrace{\int_0^\infty (\Gamma_{wt} + \Gamma_{et}) \log \ell_t - \Gamma_{wt} \frac{\ell_t^{1+1/\psi}}{1+1/\psi} dt}^{\text{labor}} + \overbrace{\int_0^\infty (\Gamma_{wt} + \Gamma_{et}) \log Z_t dt}^{\text{TFP}} \tag{H.1}$$

$$+ \underbrace{\int_0^\infty \Gamma_{wt} \log \eta_{wt} + \Gamma_{et} \left(\int p(z) \log \frac{\eta_t(z)}{p(z)} dz \right) dt}_{\text{workers vs. entrepreneurs}} - \underbrace{\int_0^\infty \Gamma_{et} \frac{1}{\rho_e} \frac{1}{2} \left(\int p(z) \sigma_{ct}^2(z) dz \right) dt}_{\text{entrepreneur risk}} \tag{H.2}$$

In principle, evaluating equation (H.2) depends on the distribution of consumption shares across entrepreneurs. In practice, we always consider policies in which the planner cannot alter the initial distribution of wealth between entrepreneurs with the same type x, nor can it alter aggregate TFP.

Then, let us denote by W^{CE}, W^{PL} total welfare in the competitive equilibrium and at the planner's allocation. We evaluate the welfare gains of moving from W^{CE} to W^{PL} in consumption-equivalent term. The welfare gain of moving from W^{CE} to W^{PL} is equivalent to increasing the consumption of entrepreneurs and workers proportionally by x percentage points, where x solves equation (H.3).

$$W^{PL} = \int_0^\infty e^{-\rho_w t}\tilde{\Gamma}_w \left(\log((1+x)c_{wt}^{CE}) - \frac{(\ell_t^{CE})^{1+1/\psi}}{1+1/\psi}\right) dt + \int_0^1 \mathbb{E}\left[\int_0^\infty e^{-\rho_e t}\tilde{\Gamma}_e \log((1+x)c_{it}^{CE}) dt\right] di \tag{H.3}$$

H.3 Welfare Gains and Transition Dynamics

In this section we evaluate welfare gains along the transition path for different policies. In the paper we computed the long-run welfare gains from two policies: the optimal labor tax and a labor subsidy equal to the average markup. In the long-run, these welfare gains depend on the average markup, the ratio between workers' consumption and their labor income, and the Frisch elasticity. However, these results on welfare gains abstract from the We use the calibrated model to evaluate welfare gains along the transition path and we find they are similar to their long-run counterparts.

Welfare gains along the transition path depend on the values of the Pareto weights Γ_{wt} and Γ_{et}. For illustrative purposes, we set $\tilde{\Gamma}_e = 0.04$ and $\tilde{\Gamma}_w = 0.96$, motivated by evidence that approximately 4% of US households are entrepreneurs (Salgado, 2020). We assume that the economy is initially in the steady state of the zero-tax equilibrium. Results along the transition path differ from the long run because the discounted Pareto weights change over time. This has two implications: the evaluated welfare gains from any policy differ from their long-run counterparts, and the optimal policy along the transition path is not necessarily equal to its long-run counterpart.

First, we report in Table H.2 welfare gains along the transition path from implementing (i) a constant labor subsidy equal to the average steady-state markup, (ii) a constant labor

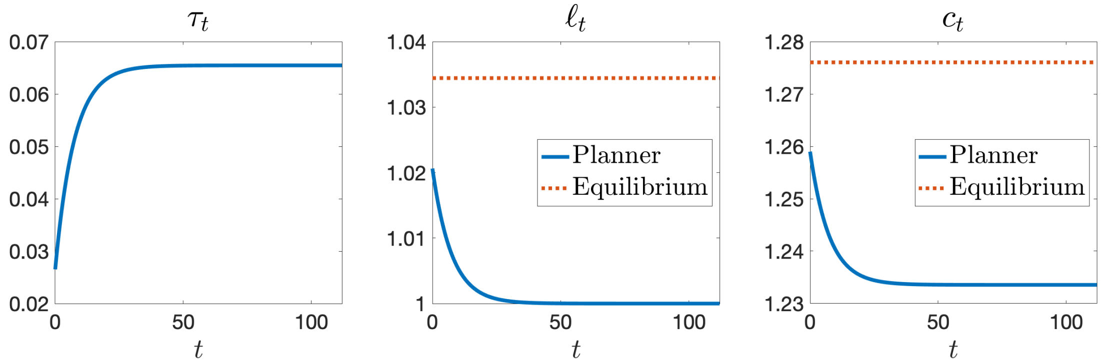


Figure H.1: Optimal labor tax τ_t, aggregate labor ℓ_t, and consumption c_t that solve the planner problem, taking as given that $t = 0$ is the steady state of the zero-tax equilibrium.

	Constant labor subsidy	Constant labor tax
	$\tau_t = \pi$	$\tau_t = \tau^*$
Welfare gains: transition path	-2.11%	0.087%
Welfare gains: long run	-2.22%	0.117%

Table H.2: Welfare Gains and Transition Dynamics.

tax equal to the optimal long-run tax. Implementing a constant labor subsidy equal to the average markup leads to substantial welfare losses equivalent to a 2.22% drop in consumption drop. As a comparison, the long-run welfare losses one would obtain using the sufficient statistic approach are equivalent to a a 2.11% drop in consumption. Thus, we find that long-run welfare gains provide a good approximation for welfare gains along the transition path.

When we evaluate welfare gains from implementing a constant labor tax equal to the optimal long-run tax, we find they are relatively small and equivalent to a 0.087% increase in consumption. As a comparison, the long-run welfare gains one would obtain using the sufficient statistic approach are equivalent to a 0.117% increase in consumption.

Finally, we implement the optimal time-varying labor subsidy. We find that welfare gains from implementing the optimal time-varying labor subsidy are similar to the welfare gains from a constant subsidy $\tau_t = \tau^*$, as they are equivalent to a 0.092% increase in consumption. Figure H.1 plots the optimal labor tax, aggregate labor, and consumption that solve the planner problem starting from the steady state of the competitive equilibrium in $t = 0$. By the separation result, the distribution of consumption shares and aggregate productivity are identical to the competitive equilibrium, and changes in welfare must come from changes in the labor term of the transformed planner's objective. The key difference from the long run is that the planner values entrepreneurs' consumption on the transition. This implies that the optimal labor tax at $t = 0$ is less than its long run value, temporarily increasing aggregate labor and consumption.

To summarize, numerical results from the transition dynamics suggest that long-run results are a good guide to optimal policy on the transition. If measured markups are compensation for risk, then mistakenly implementing the a subsidy to undo markups associated with optimal policy in the market-power perspective is very costly.