Entrepreneurship, Agency Frictions and Redistributive Capital Taxation

Corina Boar Matthew Knowles[*†]

November 11, 2020

Abstract

We study optimal capital taxation in a model with financial frictions, where the distribution of wealth across entrepreneurs of different productivity levels affects how efficiently capital is used in the economy. The planner chooses linear taxes on wealth, capital and labor income to maximize the steady state utility of a newborn agent. Most agents in the model are workers who are relatively poor, leading to a redistributive motive for capital income and wealth taxation. In our setting, optimal tax rates can be written as a closed-form function of the size of the tax bases of taxes and the elasticities of the tax base with respect to tax rates. We find that it is optimal to tax capital income at a positive rate, because general equilibrium effects on the distribution of capital across entrepreneurs attenuate the negative effects of these taxes on output. Optimal wealth taxes are close to zero, since such taxes strongly discourage capital accumulation. The tighter financial frictions are, the lower the optimal tax rate on capital income, as the planner lowers this tax in order to counteract the negative effects of financial frictions on the level and allocation of the capital stock.

1 Introduction

This paper studies optimal redistributive capital taxation in a model where taxation affects how efficiently capital is allocated in the economy. The vast literature on optimal capital

[*]We thank Teresa Steininger for superb research assistance and our discussants Nils Grevenbrock and Miguel-Angel López-García for insightful feedback. For many helpful comments, we thank Mark Bils, Yena Park, and seminar participants at the University of Leicester, University of St Andrews, University of Cologne, SUNY Binghamton, University of Barcelona, University of Kent, University of Manchester Economic Theory Workshop 2018, and the London Workshop in Quantitative Macroeconomics 2019.

[†]Corina Boar: New York University, email: corina.boar@gmail.com; Matthew Knowles: University of Cologne, email: knowles@wiso.uni-koeln.de

taxation in general equilibrium has typically analyzed models in which all physical capital is the same and the principal deadweight loss associated with capital taxation is its negative effect on aggregate saving. However, critics of high rates of capital taxation have long expressed concerns that it has harmful effects not only on the total level of investment in the economy, but also on the allocation of investment. For instance, Hayek (1960, chap. 20) argues that the taxation of profits hinders the accumulation of wealth by entrepreneurs who manage "successful new ventures", preventing these entrepreneurs from investing further. As such, he argues that "the taxation of... profits, at [high] rates, amounts to a heavy tax on that turnover of capital [between entrepreneurs] which is part of the driving force of a progressive society." Relatedly, it is often argued that taxation may affect incentives for entrepreneurs to take risks, implying that taxation may affect the allocation of capital between more and less risky uses.[1]

We analyze optimal linear capital taxation in a model which incorporates these issues. In the model, there are overlapping generations of two types of households: workers and entrepreneurs. Newborn households decide whether to become workers or entrepreneurs, and retain the same job for their entire life. A utilitarian planner sets tax rates to maximize expected utility of newborn households in the steady state and has some motivation to tax capital to redistribute from rich entrepreneurs to the poor workers. In the model, entrepreneurs choose how much capital to allocate to a risky technology and how much to allocate to a risk-free technology. Furthermore, entrepreneurs are heterogeneous in their productivity levels and lend to one another through financial markets, but these markets are frictional due to an asymmetric information problem which we model explicitly. In particular, it is assumed that financial contracts must be written to encourage entrepreneurs to truthfully report the value of their idiosyncratic shocks to productivity, rather than to lie and divert funds to themselves (in a similar vein to e.g. Bernanke, Gertler and Gilchrist, 1999). The effect of the financial friction is that entrepreneurs are limited in their ability to borrow and are unable to fully diversify idiosyncratic risks. This discourages them from allocating capital to the risky technology, which consequently has a higher expected return than the risk-free technology in equilibrium.

Together, these modeling assumptions imply that taxes on capital affect how efficiently capital is allocated in the economy. Taxes affect how much capital entrepreneurs allocate to the risky and risk-free technologies. Furthermore, the financial friction limits entrepreneurs' abilities to borrow, implying that the amount of capital used by an entrepreneur is tied to her individual wealth. Since taxes affect the distribution of wealth, they therefore affect how far capital is allocated to high productivity and low productivity entrepreneurs.

In our model, the planner chooses linear tax rates on capital income, wealth and labor

[1] See Cullen and Gordon (2007) and Devereux (2009) and the citations therein.

income. Capital income taxes (i.e. taxes on the return to capital) are not equivalent to wealth taxes (i.e. taxes on the stock of capital) in our setting, unlike in traditional models. This is because entrepreneurs with different productivity levels differ in their rates of return to capital. As such, our model speaks to recent debates about whether wealth should be taxed in addition to capital income (e.g. Saez and Zucman, 2019, Smith, Zidar and Zwick, 2020). Our model is highly tractable and we can characterize the steady state values of aggregate variables as closed form functions of prices, taxes and parameters.

We study comparative statics of the model in partial and general equilibrium. In partial equilibrium, for given prices, parameters and level of aggregate capital stock, we show that the degree to which capital is allocated to the high return risky technology is strictly decreasing in the rate of tax on capital income. As such, an increase in this tax leads to lower output for a given stock of capital and labor. Higher capital income taxes lead to a more inefficient allocation of capital because these taxes fall most heavily on entrepreneurs who have a high rate of return to capital. As such, capital income taxes tend to reduce the share of wealth of high productivity entrepreneurs, which also shifts the allocation of productive capital away from these entrepreneurs (due to financial frictions).[2] Additionally, higher taxes on capital income reduce the post-tax excess return to the risky technology, and so encourage entrepreneurs to shift capital towards the lower return risk-free technology. As such, capital income taxes shift capital away from high productivity entrepreneurs, and away from the high return risky technology, and so lead to a more inefficient allocation of capital in the steady state.

In general equilibrium, however, we find that this effect is substantially weakened. This is because, while an increase in capital income taxes leads to a reallocation of capital away from the high return risky technology, this reallocation in turn increases the pre-tax rate of return to the risky technology, which benefits the most high productivity entrepreneurs and induces them to allocate capital to the risky technology. Thus, the general equilibrium effect substantially mitigates the partial equilibrium effect.

Furthermore, we find that the steady state capital stock is typically lower when taxes on capital income and wealth are higher. This is primarily a consequence of the usual mechanism that taxes on capital discourage saving and capital accumulation. However, for capital income taxes, there are two additional effects on capital accumulation. First, higher taxes reduce the efficiency of capital allocation, as mentioned above, which reduces the average return to capital and discourages capital accumulation. Second, higher capital income taxes discourage newborn households from becoming entrepreneurs. Since entrepreneurs have decreasing returns to scale in our model, this reduces the rate of return to capital and also discourages capital accumulation. In general equilibrium, the effect of capital income taxes

[2] These effects are also operative in Guvenen et al. (2019).

on total capital accumulation is typically more moderate than the effect of wealth taxes, because the reallocative effects of capital income taxes increase the pre-tax return to the risky technology.

We show that optimal tax rates in the model can be written as a closed form function of the size of the tax base for the various taxes, the degree to which each tax is born by workers and entrepreneurs, and the elasticity of the tax bases with respect to each tax. This result is in the same vein as the literature on the "sufficient statistics" approach to optimal taxation (e.g. Piketty and Saez, 2013). Our optimal tax formula does not itself depend on many details of the model, such as functional form assumptions, or the specific details of the financial friction. However, our specific modelling assumptions allow us to determine the value of the elasticities of the tax bases with respect to tax rates.

We calibrate the model and study how optimal taxes depend on the degree of financial frictions. We find that, regardless of the severity of financial frictions, the optimal capital income tax is positive (albeit less than the optimal labor income tax) and the optimal wealth tax is negative but close to zero. In our benchmark calibration, the optimal capital income tax is 11.14% , the optimal labor income tax is 20.7%, and the optimal wealth tax is -0.44%. Optimal capital income taxes are positive, in contrast to well-known earlier results in the literature (Chamley, 1986; Judd, 1985), because of the general equilibrium effects of capital income taxes on the allocation of capital. Since capital income taxes increase the relative return of the risky technology, this substantially mitigates the negative effects of these taxes on capital accumulation, leading to a positive optimal capital income tax. Nevertheless, the planner optimally sets a slightly negative tax on wealth in order to counteract the negative effects of capital income tax on capital accumulation. Since the capital stock is very elastic with respect to wealth taxes, a slightly negative wealth tax significantly increases capital accumulation while costing little tax revenue.

We find that our allowing for endogenous entry into entrepreneurship tends to reduce the optimal capital income tax. If we shut down this channel by making the percentage of entrepreneurs exogenous, the optimal tax on capital income rises to 57.1%. Such high capital income taxes, however, make becoming an entrepreneur relatively undesirable. As such, with endogenous entry into entrepreneurship, the optimal tax on capital income is reduced (to 11.14%) as higher taxes would reduce the number of entrepreneurs and therefore output and capital accumulation.

Our results suggest that, in a setting with endogenous entry and financial frictions, optimal tax rates on capital income, wealth and labor income may not be so far from current practice in the United States, where wealth taxes are zero and labor and capital income tax rates are similar. We find that it is optimal to tax wealth at a rate close to zero and to tax capital income at a positive rate, albeit lower than the labor income tax rate.

Our model is quite general in a number of respects, which argues for the relevance of our results. In particular, our optimal tax formulae hold under weak assumptions about the production function, utility functions and the details of the financial frictions entrepreneurs face. Nevertheless, to maintain analytical tractability and simplicity of exposition, our model is restrictive on a number of important dimensions. We abstract away from many complex details of national tax codes and restrict attention to linear tax rates. We assume that workers do not vary in their ability levels and do not choose their labor effort, and we do not consider bequests, aging or transition dynamics. Finally, we do not consider the possibility of entrepreneurs declaring income as either capital or labor income in order to reduce tax liability. Relaxing these assumptions is left to future work. While these many considerations will doubtlessly affect the value of optimal taxes, the channels through which taxes affect the allocation of capital in this paper will presumably continue to operate.

The remainder of this paper is structured as follows. The next subsection reviews the recent related literature. Section 2 outlines the model assumptions. Section 3 derives properties of the model equilibrium and steady state and shows how the steady state is affected by tax rates. Section 4 derives formulae for the optimal tax rates and shows the values of optimal taxes in the numerical calibration. Section 5 concludes.

Related literature. This paper studies optimal taxation in an environment in which output depends on the allocation of capital across heterogeneous entrepreneurs, which is affected by taxation. In that sense, our paper is related to the work of Evans (2015), Shourideh (2014), Itskhoki and Moll (2019), Guvenen et al. (2019), Boar and Midrigan (2020), Basseto and Cui (2020). We differ from these papers by allowing for a wider range of tax instruments and studying the distinct effects of wealth and capital income taxes on the allocation of capital, by characterizing optimal taxes in closed form depending on "sufficient statistics" that not only enable us to carefully inspect the equity-efficiency tradeoff theoretically but also to provide a bridge between theory and empirics, and/or by micro-founding the financial friction thus allowing for changes in taxes to lead to changes in the tightness of financial frictions. Our numerical results for optimal taxes are closer to Boar and Midrigan (2020) than to Guvenen et al. (2019), in that we find that the optimal capital income tax rate is positive and the optimal wealth tax is approximately zero, whereas Guvenen et al. (2019) find large efficiency gains from shifting to wealth taxation due to reduced misallocation. Our roughly zero optimal wealth tax arises because the output losses from misallocation are relatively small in our model (consistent with results in a several related models, such as Midrigan and Xu (2014)) and so the potential efficiency gains from wealth taxation through this channel are also small.

Our paper is also related to the work that studies the effects of changing taxes numerically in models with entrepreneurs with heterogeneous productivity levels. Examples are Cagetti and De Nardi (2009), Kitao (2008), Rotberg and Steinberg (2019) who study the effect of changing estate, capital income and wealth taxes in related settings. We differ from this literature in several ways. First, our model is analytically tractable and we focus on analytical rather than numerical results, with the aim of making the intuition behind the key mechanisms as transparent as possible and exploring the effects of a wider range of tax policy changes. Second, our financial friction arises endogenously as a consequence of asymmetric information between entrepreneurs and financial intermediaries and, as such, our results highlight that the degree to which financial markets are frictional may itself be affected by changes in taxes and that this is of importance when considering optimal taxation.

Our paper also relates to Panousi and Reis (2014), Panousi and Reis (2019) and Phelan (2019), who study optimal taxation in the presence of idiosyncratic investment risk. In these papers, unlike our setting, entrepreneurs do not differ in their expected productivity levels and there is only one production technology, so the allocation of capital does not itself affect aggregate output.

Lastly, our paper contributes to the wider literature on optimal capital taxation, which generally focuses on the effect of capital taxation on aggregate capital accumulation, as in the work of Chamley (1986), Judd (1985), Straub and Werning (2014), Benhabib and Szőke (2019), Chen, Chen and Yang (2019), among others.[3] Related to our paper, Abo-Zaid (2014), Biljanovska (2019) and Biljanovska and Vardoulakis (2019) have explored how the results in this line of work are affected in settings with reduced-form financial frictions while maintaining the assumptions of Chamely and Judd that capital is homogeneous and there is no idiosyncratic risk.

The rest of the paper is organized as follows. Section 2 outlines the assumptions of the model. Section 3 discusses properties of the equilibrium of this model. Section 4 presents the planner''s optimization problem and the optimal tax policy that results. Section 5 concludes.

2 Model

In this section we describe our model economy and define an equilibrium. As we discuss in Section 4, our main result for optimal taxes does not depend on a number of the details of the model, but these details determine the numerical values of the optimal taxes that we obtain.

Environment We consider a discrete time, infinite-horizon economy populated by a continuum of measure one of households. In addition there is a continuum of competitive financial intermediaries, which we refer to as banks. Households are born identical and with

[3] See Chari and Kehoe (1999) for a survey.

no wealth. At birth each household chooses whether to be an entrepreneur or a worker and retains this occupation for their entire life. Entrepreneurs manage firms, while workers supply labor. Entrepreneurs use their capital to produce intermediate goods. In particular, each entrepreneur is the owner of two different investment projects: a risky project which produces 'risky' intermediate goods denoted by y_E, and a risk-free project, which produces 'risk-free' intermediate goods denoted by y_F.[4] After making intermediate goods, entrepreneurs sell these goods among themselves and use labor in combination with intermediate goods to make a final good. The government levies (possibly negative) taxes on households and funds the fixed (exogenously given) level of government spending $\overline{G}$.

Timing Each period t is divided into three sub-periods: morning, afternoon and evening. In the morning, entrepreneurs buy and sell capital amongst themselves and each entrepreneur freely divides her capital between her risky and her risk-free investment projects. In the afternoon, each entrepreneur draws an idiosyncratic shock which affects the quantity of capital in her risky project and her two projects produce intermediate goods. Entrepreneurs sell the intermediate goods they produce to one another. In the evening, entrepreneurs use intermediate goods and labor to produce the final good, which is sold to households. Households divide their resources between consumption and saving for the next period. At the end of the period, a fraction $\gamma \in (0,1)$ of households die and new households are born. Newborn households choose an occupation. Capital depreciates at rate $\delta \in (0,1)$.

Technology of Entrepreneurs At the beginning of each period t, each entrepreneur i is endowed with wealth $a_{i,t}$. In the morning, before capital is traded, newborn entrepreneurs draw a $\theta_{i,t}$ from the ergodic discrete distribution $g(\theta)$ given by

$$g(\theta) = \begin{cases} 1-\pi & \text{if } \theta = 0 \\ \pi & \text{if } \theta = 1 \\ 0 & \text{otherwise.} \end{cases}$$

We refer to θ as the entrepreneur's 'type'. At the start of each period, each continuing entrepreneur has the same θ as in the previous period with probability $1-\lambda_\theta$ and draws a new θ from the distribution $g(\theta)$ with probability λ_θ.

After allocating capital between her risky and risk-free projects in the morning, the entrepreneur draws a idiosyncratic shock $\epsilon_{i,t}$ in the afternoon, which is independent across time and across entrepreneurs. The shock $\epsilon_{i,t}$ affects the stock of capital in the entrepreneur's

[4]The device of having two separate types of intermediate goods is a simple way to allow entrepreneurs to choose between allocating capital in a risky way or risk-free way. This is designed to capture the idea that some investment projects are more risky than others and that capital owners must take into account the risks associated with different projects when making investment decisions.

risky project, so that an entrepreneur who allocates $k_{E,i,t}$ to her risky project in the morning of period t, and draws the shock $\epsilon_{i,t}$, has $\epsilon_{i,t}k_{E,i,t}$ units of capital in her risky project in the afternoon. As such, the quantity of capital in this project changes stochastically over time, rendering the project risky. We assume that each entrepreneur's $\epsilon_{i,t}$ is drawn from a lognormal distribution H. In particular, we assume that

$$\epsilon_{i,t} = \underline{\epsilon} + (1 - \underline{\epsilon}) \exp\left(\varphi\xi_{i,t} - \frac{\varphi^2}{2}\right),$$

where $\xi_{i,t} \sim N(0, 1)$, $\underline{\epsilon} \in (0, 1)$ and φ is a parameter determining the variance of ϵ. This assumption implies that $\mathbb{E}[\epsilon_{i,t}] = 1$ and $\mathrm{Var}(\log(\epsilon_{i,t} - \underline{\epsilon})) = \varphi^2$. The lowest possible realization of ϵ is $\underline{\epsilon}$.

Each unit of capital in the risky project produces risky intermediate goods equal to the entrepreneur's $\theta_{i,t}$. Therefore, if an entrepreneur with type $\theta_{i,t}$ allocates $k_{E,i,t}$ to her risky project in the morning of period t, then in the afternoon the risky project has $\epsilon_{i,t}k_{E,i,t}$ units of capital, and produces $\theta_{i,t}\epsilon_{i,t}k_{E,i,t}$ units of risky intermediate goods. After allocating capital to her risky project, the entrepreneur allocates any remaining capital $k_{F,i,t}$ to her risk-free project. In the afternoon, this project produces an output of $y_{F,i,t} = k_{F,i,t}$ risk-free intermediate goods.

In addition to producing intermediate goods, entrepreneurs are able to hide capital $k_{H,i,t}$ in their risky project after observing their shock $\epsilon_{i,t}$ and convert it directly into units of consumption.[5] In particular, instead of using units of capital in the risky project to produce intermediate goods, an entrepreneur can convert one unit of capital in the risky project into $\phi \in (0, 1)$ units of consumption. It will be shown that, when taxes are set optimally, entrepreneurs will not choose to hide any units of capital. However, the ability of entrepreneurs to hide units of capital affects allocations and optimal taxes by creating frictions in financial markets.

Technology of Final Good Production We assume that entrepreneurs trade risky intermediate goods among themselves at price $r_{E,t}$ per unit and risk-free intermediate goods at price $r_{F,t}$ per unit. Since an entrepreneur's output of the risk-free intermediate good is $y_{F,i,t} = k_{F,i,t}$, a consequence of this is that $r_{F,t}$ is the market rate of return to capital in risk-free projects per period. Each entrepreneur i hires $n_{i,t}$ workers at wage rate w_t and uses $y^d_{E,i,t}$ and $y^d_{F,i,t}$ units of risky and risk-free intermediate goods to make $y_{i,t}$ final goods according to the production function

$$y_{i,t} = f\left(y^d_{E,i,t}, y^d_{F,i,t}, n_{i,t}\right).$$

[5] As we subsequently discuss, the realization of $\epsilon_{i,t}$ is private information to the entrepreneurs, so they are able to hide capital.

We assume that f is concave and strictly increasing in all arguments, exhibits strictly decreasing returns to scale and satisfies the Inada conditions.

Preferences Each worker has a constant labor endowment $n = 1$ which he supplies inelastically. Entrepreneurs do not supply labor. The consumption of an entrepreneur i is denoted $c_{i,t}$ and that of a worker i is denoted by $c^N_{i,t}$. Households have period utility function $U(c) \equiv \log(c)$ and maximize expected lifetime utility.

Occupational choice Newborn households choose their occupation to maximize expected lifetime utility. Since newborn households are identical, the number of entrepreneurs and workers in equilibrium will adjust until newborn households at each t are indifferent between the two occupations:

$$\sum_{j=1}^{\infty} (1-\rho)^{j-1} (1-\gamma)^{j-1} \log(c^N_{i,t+j}) = \mathbb{E}_t \left[\sum_{j=1}^{\infty} (1-\rho)^{j-1} (1-\gamma)^{j-1} \log(c_{i,t+j}) \right],$$

where the expectation is with respect to the future realizations of $\theta_{i,t}$ and $\epsilon_{i,t}$. Without loss of generality, we assume that if $i \leq N_t$ the household is a worker and $i > N_t$ the household is an entrepreneur.

Government The government levies four different types of tax: a consumption tax $\tau_{C,t}$, a labor income tax $\tau_{N,t}$, a capital income tax $\tau_{K,t}$ and a wealth tax $\tau_{W,t}$, and has to finance exogenous expenditure $\overline{G}$, while balancing its budget every period. Taxes are paid in the evening and government spending also takes place in the evening. The government is not allowed to trade in financial assets at any time. The government's budget constraint each period is

$$\overline{G} = \tau_{N,t} w_t N_t + \tau_{K,t}(\Pi_t - \delta K_t) + \tau_{W,t} K_t + \tau_{C,t}(C_t - \phi K_{H,t}), \tag{1}$$

where N_t is the total measure of workers, K_t is the aggregate capital stock at the start of the period and $\Pi_t - \delta K_t$ is the total reported profits of entrepreneurs net of capital depreciation. C_t is aggregate consumption and $\phi K_{H,t}$ is the total consumption entrepreneurs generate by hiding capital, which cannot be taxed. We show below that $K_{H,t}$ is always zero in the steady state if the tax policy is set optimally.

Financial Markets Entrepreneurs may fund capital purchases each morning by writing one-period state-contingent financial contracts with banks. We assume that banks are risk neutral and perfectly competitive and live for only one period each, so they have no interest in multi-period financial contracts. New banks are created at the start of each period. The financial market opens immediately after each entrepreneur's ability $\theta_{i,t}$ is revealed. If she writes a financial contract with the bank, the entrepreneur receives from the bank some quantity $b_{i,t}$ (possibly negative) in the morning and in exchange she agrees to return to the bank the quantity $\hat{b}_{i,t}$ (possibly negative) at the end of the period, where $\hat{b}_{i,t}$ may depend on the realization of the entrepreneur's shock $\epsilon_{i,t}$. In this way, financial contracts function as a within-period loan for entrepreneurs, and entrepreneurs can also use them to insure themselves against the idiosyncratic risk associated with the shock $\epsilon_{i,t}$. We refer to an entrepreneur as a borrower if she chooses $b_{i,t} > 0$ and a saver if she chooses $b_{i,t} < 0$.

It is convenient to write the entrepreneur's choices of $b_{i,t}$ and $\hat{b}_{i,t}$ as policy functions of the relevant state variables. In general, an entrepreneur's choice of $b_{i,t}$ will depend on her ability $\theta_{i,t}$, her start of period capital $k_{i,t}$ and the aggregate state of the economy, which we label X_t. Therefore, abusing notation slightly, we write $b_{i,t} \equiv b(a, \theta, X)$. Likewise, we write $\hat{b}_{i,t} \equiv \hat{b}(a, \theta, \epsilon, X)$, since $\hat{b}_{i,t}$ will depend on the entrepreneur's individual states a, θ, ϵ and the aggregate state X.

Since banks are risk-neutral, perfectly competitive and profit maximizing, a bank will agree to a financial contract written by an entrepreneur if and only if the financial contract delivers it non-negative profits in expectation at the end of the period. As such, banks will only lend to entrepreneurs in the morning if the expected return on the loan in the evening is equal to the market risk-free rate. This implies the following constraint

$$\int_{\epsilon} \hat{b}\left(a, \theta, \epsilon, X\right) \mathrm{d}H\left(\epsilon\right) \geq R_{F,t} b\left(a, \theta, X\right),$$

where $R_{F,t}$ denotes the gross market risk-free rate of interest within the period. Since entrepreneurs have no desire to pay the banks more than is necessary or accept an interest rate less than $R_{F,t}$ if they are lending to the banks, this inequality will be satisfied with equality. The consequence is that banks make zero profits in equilibrium.[6]

Workers can also borrow or lend to banks within the period at the risk free rate $R_{F,t}$.

Annuities At the end of the period households may trade among themselves in financial annuities, which insure against the risk of death. A household may exchange 1 unit of the good at the end of the period for $\frac{1}{1-\gamma}$ units of the good at the start of the next period if the household is still alive (or the converse). Entrepreneurs place all their capital in a common fund at the end of the period, exchanging it for annuities.[7]

[6]Since banks make zero profits, it makes no difference to the equilibrium behavior of the economy who owns the banks. We may assume that they are owned either by workers or by entrepreneurs.

[7]If we allowed entrepreneurs to hold capital rather than placing in the common fund, they would prefer to place in the common fund.

Budget Constraints Let a worker i's start of period assets be denoted $a^N_{i,t}$. The worker's asset position evolves according to:[8]

$$c^N_{i,t}(1+\tau_{C,t}) + (1-\gamma)a^N_{i,t+1} \quad = \quad w_t(1-\tau_{N,t}) + R_{F,t}a^N_{i,t}.$$

In the morning of each period the entrepreneur may buy and sell capital, divide her capital between a risky and risk-free project and write a financial contract with a bank. Her choices in the morning must satisfy the budget constraint

$$k_{E,i,t} + k_{F,i,t} = k_{i,t} = a_{i,t} + b_{i,t}.$$

After receiving the $\epsilon_{i,t}$ shock, the entrepreneur chooses how many units (if any) of capital in the risky project to hide. We let $k_{H,i,t}$ denote the quantity of capital the entrepreneur hides in the afternoon. In the evening, the entrepreneur chooses how much to consume, which we denote by $c_{i,t}$, and how many annuities to buy, $(1-\gamma)a_{i,t+1}$. Capital hidden in the afternoon is transformed into $c_{H,i,t}$ units of consumption. Finally, in the evening the entrepreneur repays the bank $\hat{b}_{i,t}$ (or is paid by the bank if $\hat{b}_{i,t} < 0$) and pays her taxes to the government. Consequently, in the evening the entrepreneur's budget constraint is

$$c_{i,t} - c_{H,i,t} + (1-\gamma)a_{i,t+1} + \hat{b}_{i,t} \quad = \quad \pi_{i,t} - T_{i,t} + (1-\delta)k_{i,t},$$

where $\pi_{i,t}$ is the entrepreneur's period profits given by

$$\begin{aligned} \pi_{i,t} \quad &= \quad \underbrace{(r_{E,t}y_{E,i,t} + r_{F,t}y_{F,i,t})}_{\text{profit from intermediate goods}} + \underbrace{\left(y_{i,t} - w_t n_{i,t} - r_{E,t}y^d_{E,i,t} - r_{F,t}y^d_{F,i,t}\right)}_{\text{profit from final good}} \\ &+ \quad \underbrace{(1-\delta)\left((\epsilon_{i,t}-1)\,k_{E,i,t} - k_{H,i,t}\right)}_{\text{reported capital gains}}, \end{aligned}$$

$T_{i,t}$ is the entrepreneur's period tax payments given by

$$T_{i,t} = \tau_{C,t}\left(c_{i,t} - c_{H,i,t}\right) + \tau_{K,t}\pi_{i,t} - \tau_{K,t}\delta k_{i,t} + \tau_{W,t}k_{i,t},$$

[8]This budget constraint indicates that workers do not pay capital income or wealth taxes, which may seem strange. For analytical convenience, we assume that wealth and capital income taxes are levied on physical assets and entrepreneurial profits only, rather than on financial assets. As such, workers do not pay capital income and wealth taxes. This assumption is without loss of generality. If linear tax rates were also imposed on individual's net financial positions and net financial wealth, these taxes would simply lead to an adjustment of pre-tax interest rates, and leave the post-tax rate of interest and allocations unaffected, as in the classic discussion in Varian (2014), p. 307.

and where $c_{H,i,t}$, $y_{E,i,t}$, $y_{F,i,t}$ satisfy

$$\begin{aligned} c_{H,i,t} &= \phi k_{H,i,t} \\ y_{E,i,t} &= \theta_{i,t}\epsilon_{i,t}k_{E,i,t} \\ y_{F,i,t} &= k_{F,i,t}. \end{aligned}$$

Agency Friction During the period, an entrepreneur's realization of ϵ, the quantity of capital in the risky sector she hides and the consumption she obtains from converting hidden capital are all private information. In particular, after observing the shock ϵ, an entrepreneur can choose to honestly report the amount of capital she has in the risky sector, but she can also lie by under-reporting the amount of capital she has and hiding more capital than she admits to. However, the quantity of capital allocated to the entrepreneur's projects initially, and the quantity of intermediate goods she produces are assumed to be public information.[9]

When an entrepreneur writes a financial contract in the morning, the market will expect the entrepreneur to repay $\hat{b}(a,\theta,\epsilon,X)$ in the evening, given her realization of ϵ. In equilibrium, the market must be correct in expecting this, and so it is without loss of generality to restrict attention to incentive compatible contracts where the entrepreneur honestly reports her ϵ, and pays the promised amount $\hat{b}(a,\theta,\epsilon,X)$. The entrepreneur will be tempted to lie about ϵ only if doing so increases her available resources for consumption and/or her next period capital. This gives rise to the following incentive compatibility constraint

$$\frac{(1-\tau_K)\left(r_E+(1-\delta)\right)k_E}{1+\tau_C} \geq \phi k_E + \frac{\partial \hat{b}(a,\theta,\epsilon,X)}{\partial \epsilon}\frac{1}{1+\tau_C},$$

The left-hand-side of this constraint represents the cost of under-reporting ϵ expressed in units of consumption and the right-hand-side represents the benefit of under-reporting ϵ. Specifically, by under-reporting ϵ by some small amount $\mathrm{d}\epsilon$ the entrepreneur will find herself with $k_E \mathrm{d}\epsilon$ fewer units of capital and loses the after-tax return on those. At the same time she transforms the $k_E \mathrm{d}\epsilon$ hidden units of capital into $\phi k_E \mathrm{d}\epsilon$ units of consumption and also repays less to the bank according to the optimal financial contract.

Worker's Optimization Problem The worker's problem is to choose a level of consumption each period to maximize present discounted utility. The worker's state variables are his assets a^N and the aggregate state X. As such, the worker's consumption and next period assets are given by the policy functions $c^N(a^N,X)$ and $a^{N\prime}(a^N,X)$ that solve the

[9] In the extreme case $\phi = 0$ there would be no informational friction, since the entrepreneur has no incentive to hide capital.

following Bellman equation:

$$V^N(a^N, X) = \max\ \log\left(c^N\right) + (1-\rho)\,(1-\gamma)\,V^N\left(a^{N\prime}, X'\right)$$

subject to the worker's budget constraint.

Entrepreneur's Optimization Problem The entrepreneur's optimization problem is to choose functions $k_E(\cdot) \geq 0$, $k_F(\cdot) \geq 0$, $k_H(\cdot) \geq 0$, $b(\cdot)$, $\hat{b}(\cdot)$, $c(\cdot) \geq 0$, $c_H(\cdot) \geq 0$, $y_E(\cdot) \geq 0$, $y_F(\cdot) \geq 0$, $y_E^d(\cdot) \geq 0$, $y_F^d(\cdot) \geq 0$, $n(\cdot) \geq 0$, $y(\cdot) \geq 0$ and $a'(\cdot)$ to solve

$$\begin{aligned} V(a,\theta,X) &= \sup \int_{\epsilon>0} \Bigg(\log(c(a,\theta,\epsilon,X)) \\ &+ (1-\rho)(1-\gamma)E\left[V(a'(a,\theta,\epsilon,X),\theta',X')\middle|\theta\right]\Bigg)\mathrm{d}H(\epsilon), \end{aligned}$$

subject to the budget constraints in the morning and in the evening, the production functions for c_H, y_E, y_F and y, the incentive compatibility constraint and the break-even condition for the banks.

Here, by having the entrepreneur choose the functions $b(\cdot)$ and $\hat{b}(\cdot)$ subject to the incentive compatibility constraint and break-even condition for the bank, we are assuming that the entrepreneur designs a financial contract and proposes it to a bank. The bank accepts provided that the contract is incentive compatible and the bank breaks even in expectation.

Aggregation and Market Clearing The aggregate level of consumption C_t and of $C_{H,t}$ satisfy

$$\begin{aligned} C_t &= \int_{i \leq N_t} c_{i,t}^N \mathrm{d}i + \int_{i>N_t} c_{i,t}\mathrm{d}i \\ C_{H,t} &= \int_{i>N_t} c_{H,i,t}\mathrm{d}i \leq C_t. \end{aligned}$$

The aggregate levels of capital devoted to each use and the aggregate level of final good output sold are given by

$$
\begin{aligned}
K_{E,t} &= \int_{i>N_t} k_{E,i,t}\mathrm{d}i \\
K_{F,t} &= \int_{i>N_t} k_{F,i,t}\mathrm{d}i \\
K_t &= K_{E,t} + K_{F,t} \\
K_{H,t} &= \int_{i>N_t} k_{H,i,t}\mathrm{d}i \\
Y_t &= \int_{i>N_t} y_{i,t}\mathrm{d}i.
\end{aligned}
$$

Total reported period profits of entrepreneurs are

$$
\Pi_t = Y_t - w_t N_t - (1-\delta) K_{H,t},
$$

where $-(1-\delta) K_{H,t}$ is the average reported capital gain. In each period, the asset market must clear. This requires that the total capital stock equals the wealth of entrepreneurs and workers

$$
\int_{i\leq N_t} a_{i,t}^N \mathrm{d}i + \int_{i>N_t} a_{i,t}\mathrm{d}i = K_t
$$

The market for intermediate goods of each type must clear each period

$$
\begin{aligned}
\int_{i>N_t} y_{E,i,t}^d \mathrm{d}i &= \int_{i>N_t} y_{E,i,t}\mathrm{d}i \\
\int_{i>N_t} y_{F,i,t}^d \mathrm{d}i &= \int_{i>N_t} y_{F,i,t}\mathrm{d}i.
\end{aligned}
$$

The labor market must clear each period

$$
\int_{i>N_t} n_{i,t}\mathrm{d}i = N_t
$$

The final goods market clearing condition then follows by Walras' law

$$
\bar{G} + C_t + K_{t+1} = Y_t + (1-\delta) K_t - (1-\delta) K_{H,t} + C_{H,t}.
$$

Equilibrium We are now in the position to define an equilibrium for our model economy.

Definition 1. *Given a sequence of tax rates $\{\tau_{W,t}, \tau_{K,t}, \tau_{C,t}, \tau_{N,t}\}_{t=0}^{\infty}$, an equilibrium $\mathcal{E}$ is a sequence of prices $\{R_{F,t}, r_{E,t}, r_{F,t}, w_t\}_{t=0}^{\infty}$, policy functions giving entrepreneurs' and workers' decisions and a sequence of aggregate variables $\{C_t, C_{H,t}, K_t, K_{E,t}, K_{F,t}, K_{H,t}, Y_t, N_t\}_{t=0}^{\infty}$ such* *that:*

1. *The government's budget constraint is balanced every period.*
2. *Workers' decision rules solve the worker's optimization problem.*
3. *Entrepreneurs' decision rules are given by the solution to the entrepreneur's problem.*
4. $\{C_t, C_{H,t}, K_t, K_{E,t}, K_{F,t}, K_{H,t}, Y_t\}_{t=0}^{\infty}$ *represent the aggregate of household's decisions defined above.*
5. *Newborn agents are indifferent between being entrepreneurs and workers.*
6. *The asset, intermediate goods and labor markets clear.*

3 Properties of the Model Equilibrium

In this section, we characterize the equilibrium of the economy. We do so by characterizing the optimal decisions of households and the aggregate steady state of the model. We derive comparative static results for how aggregate steady state variables change in response to changes in taxes which we then use in characterizing optimal taxes.

3.1 Worker's Optimal Decisions

We first solve the worker's optimization problem. Let P_t^N denote the discounted value of a worker's lifetime income. The expression for P_t^N is given by

$$P_t^N = a_t^N + \underbrace{\sum_{j=0}^{\infty} \left[\frac{w_{t+j}(1-\tau_{N,t+j})(1-\gamma)^j}{\Pi_{k=0}^{j} R_{F,t+k}} \right]}_{F_t^N}.$$

Note that the worker's problem can be reformulated as one in which the worker decides between today's consumption c_t^N and tomorrow's lifetime wealth P_{t+1}^N, based on today's lifetime wealth P_t^N and the aggregate state X.[10] In Appendix A.1, we show that the solution is given by

$$c_t^N = \left[1 - (1-\rho)(1-\gamma)\right] \frac{R_{F,t}}{1+\tau_{C,t}} P_t^N,$$
$$P_{t+1}^N = (1-\rho) R_{F,t} P_t^N,$$

[10]This follows from a^N and P^N being deterministically linked through the expression for P_t^N.

and that the associated value function is

$$V^N(P^N, X) = V^N(1, X) + \frac{1}{1-(1-\rho)(1-\gamma)} \log P^N.$$

The worker hence devotes fraction $1-(1-\rho)(1-\gamma)$ of her discounted lifetime income (evaluated at the end of the period after she has accrued interest R_F) to consumption expenditure.

3.2 Entrepreneur's Optimal Decisions

We now solve the entrepreneur's optimization problem. To simplify the problem, note first that all the entrepreneur ultimately cares about this period is her level of consumption c and her level of wealth for the next period, a'. These are the only variables over which she has influence that enter directly into the entrepreneurs' Bellman equation. The levels of c and a' that the entrepreneur can afford at the end of the period depend solely on her total resources and the end of the period. That is, her within-period choices of how much capital to put into each project, how much to borrow and how much capital to hide only depend on the level of c and a' she can afford insofar as they affect the resources she will have at the end of the period. Therefore, we can split the entrepreneur's problem into a within-period choice of trying to achieve a high value of end-of-period resources, and a between period choice of how to divide her resources between consumption and next-period wealth.

To this end, consider first the entrepreneur's optimal production of final goods. This is a static problem in which the entrepreneur chooses intermediate inputs $y^d_{E,i,t}$, $y^d_{F,i,t}$ and labor $n_{i,t}$ to maximize profits

$$f\left(y^d_{E,i,t}, y^d_{F,i,t}, n_{i,t}\right) - r_{E,t} y^d_{E,i,t} - r_{F,t} y^d_{F,i,t} - w_t n_{i,t}.$$

Since $f(\cdot)$ exhibits decreasing returns to scale, at the optimal choice of $y^d_{E,i,t}, y^d_{F,i,t}, n_{i,t}$, the entrepreneur makes profits $\pi^\star_t$ per period from final goods production, which is the same across entrepreneurs.

Let $P_{i,t}$ denote the present value of lifetime resources that an entrepreneur i could obtain if she never produces intermediate goods and lends lends her endowment to banks at the risk-free rate $R_{F,t}$; $P_{i,t}$ is equal to

$$P_{i,t} = a_{i,t} + \underbrace{\sum_{j=0}^{\infty} \left[\frac{\pi^\star_{t+j}(1-\tau_{K,t+j})(1-\gamma)^j}{\Pi_{k=0}^{j} R_{F,t+k}} \right]}_{F_t}.$$

Then, the entrepreneur's evening budget constraint can be re-written as

$$c_{i,t} + \frac{(1-\gamma)\,P_{i,t+1}}{(1+\tau_{C,t})} = \omega_{i,t},$$

where $\omega_{i,t}$ satisfies

$$\omega_{i,t} = \frac{(1+\tau_{C,t})\,c_{H,i,t} - \hat{b}_{i,t} + (1-\tau_{K,t})\,\pi_{i,t} + (\tau_{K,t}\delta - \tau_{W,t})\,k_{i,t} + (1-\delta)k_{i,t} + (1-\gamma)\,F_{t+1}}{(1+\tau_{C,t})}.$$

We refer to $\omega_{i,t}$ as end-of-period lifetime resources.

To split the entrepreneur's problem into a within and a between-period choice, let $\tilde{V}(\omega, X)$ denote the value in the evening of a period of an entrepreneur with lifetime resources ω, who is yet to divide her resources between c and P'. Then, abusing notation, we can write the entrepreneur's between period problem recursively as:[11]

$$\tilde{V}(\omega, X) = \sup_{c,P'} \left(\log(c) + (1-\rho)\,(1-\gamma)\mathbb{E}V(P', \theta', X') \right), \tag{2}$$

$$\text{s.t.} \quad c + \frac{(1-\gamma)\,P'}{1+\tau_C} = \omega. \tag{3}$$

The entrepreneur's recursive within-period problem is to choose non-negative functions $k_E(P,\theta,X)$, $k_F(P,\theta,X)$, $k_H(P,\theta,X)$, $\omega(P,\theta,\epsilon,X)$ and functions $b(P,\theta,X)$, $\hat{b}\,(P,\theta,\epsilon,X)$ to solve:

$$V(P,\theta,X) = \sup \int_{\epsilon>0} \tilde{V}(\omega, X)\mathrm{d}H(\epsilon),$$

subject to the constraints:

$$\begin{aligned}
k_E + k_F &= a + b = P + b - F && (4)\\
\int_\epsilon \hat{b}\mathrm{d}H\,(\epsilon) &= R_F b && (5)\\
(1+\tau_C)\,\omega &= (1+\tau_C)\,\phi k_H - \hat{b} + (1-\tau_K)\,\pi + (\tau_K\delta - \tau_W)\,k + (1-\delta)k + (1-\gamma)\,F' && (6)\\
F &= \frac{\pi^\star\,(1-\tau_K)}{R_F} + F'\frac{1-\gamma}{R_F} && (7)\\
\frac{\partial\omega}{\partial\epsilon} &\geq \phi k_E, && (8)
\end{aligned}$$

where the last constraint is the incentive compatibility constraint obtained after subsituting in the definition of ω in (6).

The constant returns to scale assumptions on the entrepreneur's technology for producing intermediate goods means that the value function must take a particular form, as shown in

[11] We can relabel $V\,(a,\theta,X)$ as $V\,(P,\theta,X)$ since $P = a + F$ and F depends only on the aggregate state X.

the following lemma. This considerably simplifies the solution to the entrepreneur's problem.

Lemma 1. *There exists a function* $\overline{V}(\theta, X)$ *such that, for any* P, θ *and* X,

$$V(P, \theta, X) = \overline{V}(\theta, X) + \frac{\log(P)}{1 - (1-\rho)(1-\gamma)},$$

where $\overline{V}(\theta, X) = V(1, \theta, X)$.

Proof. See Appendix A.2. □

Using Lemma 1, the solution to the between period problem can be found immediately by taking the first order condition

$$\frac{1}{(1+\tau_C)\omega - (1-\gamma)\, P'} = \frac{1-\rho}{1-(1-\rho)(1-\gamma)}\frac{1}{P'},$$

and combining it with equation (3) to conclude that the entrepreneur chooses

$$\begin{aligned} c &= (1 - (1-\rho)(1-\gamma))\,\omega \\ P' &= (1+\tau_C)(1-\rho)\,\omega. \end{aligned}$$

Substituting these choices into the Bellman equation (2), we have that

$$\begin{aligned} \tilde{V}(\omega, X) = \frac{\log(\omega)}{1-(1-\rho)(1-\gamma)} &+ \log(1-(1-\rho)(1-\gamma)) \\ &+ \frac{(1-\rho)(1-\gamma)\log((1+\tau_C)(1-\rho))}{1-(1-\rho)(1-\gamma)} \\ &+ (1-\rho)(1-\gamma)\,\mathbb{E}\left[\overline{V}(\theta', X')\right]. \end{aligned}$$

This completes the solution of the between period problem.

To solve the within period problem, we note first note that, since the entrepreneur is risk averse and the bank is risk neutral, she will want to choose a contract that minimizes the variance of ω while observing the incentive compatibility constraint (8). Hence, this constraint must bind with equality. Integrating the resulting incentive compatibility constraint with respect to ϵ, it follows that there must exist some function $\underline{\omega}(P, \theta, X)$ such that:

$$\omega(P, \theta, \epsilon, X) \equiv \underline{\omega}(P, \theta, X) + \phi k_E(P, \theta, X)\epsilon. \tag{9}$$

In the absence of agency frictions, the entrepreneur and the bank would prefer a contract in which the bank took all the risk and the entrepreneur's ω was independent of ϵ. The agency friction prevents this, leading the entrepreneur to face the level of risk implied by the equation (9).

To simplify the within period problem further, note that during the period t the entrepreneur can borrow risk-free capital at the bank at gross interest rate of R_F. This stems from the lack of informational asymmetry in the risk-free project and the bank's zero profit condition. Moreover, the entrepreneur may sell the output from the risk-free project to other entrepreneurs to be used in the production of the final good. Hence, to rule out the presence of arbitrage, the two returns have to be equalized:

$$R_F = 1 + \left[(1 - \tau_K) (r_F - \delta) - \tau_W \right]. \tag{10}$$

The consequence is that in equilibrium, the entrepreneur will weakly prefer to lend to the bank over investing in the risk-free project.

Similarly, the entrepreneur has two different risky activities she can undertake: producing risky intermediate goods, or hiding capital after observing her ϵ shock. In equilibrium, it must be the case that the post-tax return to producing risky intermediate goods is weakly higher than the return to hiding capital. If not, all entrepreneurs would hide all their capital k_E and produce no risky intermediate goods and the Inada conditions on the final goods production function would drive the return r_E to infinity. Therefore entrepreneurs weakly prefer to use their capital to produce risky intermediate goods rather than hiding it.

Given that the entrepreneur weakly prefers to lend to the bank to producing risk-free intermediate goods and weakly prefers to produce risky intermediate goods rather than hiding capital, we can solve her optimization problem under the assumption that she chooses $k_F = k_H = 0$, with the understanding that in equilibrium some entrepreneurs may choose $k_F > 0$ or $k_H > 0$, to the extent needed to clear markets. Combining the rewritten incentive compatibility constraint (9) with the definition of ω, F, profits π, integrating with respect to ϵ and using that $k_E + k_F = k$, $k_F = k_H = c_H = 0$ and $\mathbb{E}[\epsilon] = 1$, reveals that $\underline{\omega}(\cdot)$ must satisfy

$$(1 + \tau_C)(\underline{\omega} + \phi k_E) = (1 - \delta(1 - \tau_K) - \tau_W) k_E + (1 - \tau_K) r_E \theta k_E - \int_\epsilon \hat{b}\,(\cdot)\, \mathrm{d}H\,(\epsilon) + R_F F.$$

Combining this with the bank zero profit condition (5) and the budget constraint (4) and rearranging gives

$$(1 + \tau_C)\underline{\omega} = (1 - \delta(1 - \tau_K) - \tau_W - \phi(1 + \tau_C)) k_E + (1 - \tau_K) r_E \theta k_E + R_F (P - k_E).$$

Using the above together with the no-arbitrage condition for risk-free returns (10), we can rewrite the entrepreneur's within-period problem more compactly. The entrepreneur seeks to choose functions $k_E(P,\theta,X) \geq 0$ and $\underline{\omega}(P,\theta,X)$ to solve:

$$\sup \int_\epsilon \log(\underline{\omega} + \phi\epsilon k_E)\mathrm{d}H(\epsilon),$$

subject to the constraints:

$$\begin{aligned}\underline{\omega} &= \left(-\phi + (r_E\theta - r_F)\frac{1-\tau_K}{1+\tau_C}\right)k_E + \left(\frac{1}{1+\tau_C}R_F\right)P \\ \underline{\omega} + \phi\underline{\epsilon}k_E &\geq 0.\end{aligned}$$

Here, we used that the only part of $\tilde{V}$ which depends on the entrepreneur's decisions is the term $\frac{\log(\omega)}{1-(1-\rho)(1-\gamma)}$. Therefore, maximizing the expected value of $\tilde{V}$ amounts to maximizing the expected value of this term. The second constraint arises because the minimum value of ϵ is $\underline{\epsilon} > 0$ and the entrepreneur can be sure of non-negative consumption if and only if $\underline{\omega} + \rho\underline{\epsilon}k_E \geq 0$.

The entrepreneur's within-period optimization problem amounts to determining the optimal choice of k_E. This is simply a problem of a trade-off between risk and return. Choosing higher k_E increases the variance of ω, since $\omega = \underline{\omega} + \phi\epsilon k_E$, but higher k_E may carry a higher expected return.

These results imply that the optimal financial contract between the entrepreneur and bank takes an easily interpretable form as an equity and debt contract, as discussed in the following Lemma.

Lemma 2. *The optimal financial contract is equivalent to a contract in which the entrepreneur takes a loan equal to fraction R_F^{-1} of the end of period value her risky project under the worst possible realization of ϵ, and sells fraction $1 - \frac{\phi(1+\tau_C)}{(1-\tau_K)(1-\delta)}$ of the value of her risky project as equity, retaining the remaining equity herself.*

Proof. See Appendix A.3 □

The reason that the entrepreneur is unable to sell all the equity in her project is that she needs to have a large enough 'skin in the game' in order to encourage her not to hide capital in her project. An important consequence of the endogenous modelling of financial constraints here is that the fraction of the value of the entrepreneur's project that she must retain as 'skin in the game' is itself dependent on taxes. Higher levels of capital income tax reduce the fraction of equity in her project that she is able to sell, thus tightening the financial frictions the entrepreneur faces. This creates a stronger motive against taxing capital than would occur if the fraction of 'skin in the game' the entrepreneur needed was fixed as an exogenous parameter.

Now, we note that, in an equilibrium of this economy, it must be that the price of risky intermediate goods is sufficiently low that entrepreneurs do not wish to produce and sell infinite quantities of risky intermediate goods. This condition is contained in the following lemma.

Lemma 3. *In an equilibrium of the economy, it must be the case every period that:*

$$-\phi + (r_E - r_F)\frac{1-\tau_K}{1+\tau_C} \quad \leq \quad 0 \tag{11}$$

Proof. See Appendix A.4. □

3.3 Continuous Time Limit

To describe the environment, the equilibrium conditions and to simplify the contracting problem between the entrepreneurs and banks it was natural to make the assumption of discrete time. In the remainder of the paper, which is devoted to characterizing the steady state of the economy and solving for the optimal taxes, it is more convenient to work in continuous time. We formally derive a continuous-time version of our discrete-time model in Appendix B. There, we assume that each period is of length Δ and obtain solutions to the entrepreneur's problem and characterize the steady state of the economy. We then take the limit as Δ goes to zero. This leads to the following optimal decision rule for the entrepreneur and worker in the continuous-time version of the model.

To simplify the optimal decision rules, let $\tilde{R}_F$ denotes the net risk-free rate of return:

$$\tilde{R}_F := R_F - 1$$

Proposition 1. *In equilibrium, the unique solution of the worker's problem is:*

$$c^N \quad = \quad (\rho+\gamma)\frac{P^N}{1+\tau_C} \tag{12}$$

$$dP^N \quad = \quad \left\{\left[\tilde{R}_F + \gamma\right] P^N - (1+\tau_C)c^N\right\} dt, \tag{13}$$

The unique solution of the entrepreneur's problem is:

$$\hat{k}_E \quad = \quad \frac{k_E}{P} = \frac{1}{(1+\tau_C)\phi(1-\underline{\epsilon})} \times \max\left\{0; \min\left\{\frac{(\theta r_E - r_F)(1-\tau_K)}{(1+\tau_C)\phi(1-\underline{\epsilon})\varphi^2}; 1\right\}\right\} \tag{14}$$

$$c \quad = \quad (\rho+\gamma)\frac{P}{1+\tau_C} \tag{15}$$

$$\begin{aligned} dP \quad &= \quad \left\{\left[\tilde{R}_F + \gamma\right] P + \left[k_E\left(\theta r_E - r_F\right)\left(1-\tau_K\right) - (1+\tau_C)c\right]\right\} dt \\ &+ \quad k_E\phi\left(1+\tau_C\right)\left(1-\underline{\epsilon}\right)\varphi dW, \end{aligned} \tag{16}$$

where dW is the difference of a standard Brownian motion.

Proof. See Appendix B.2 and Appendix B.3. □

As such, the solution to the entrepreneur's optimization problem implies that if the after-tax expected return on the risky project $\theta r_E(1-\tau_K)$ is lower than the after-tax risk-free net return on the risk-free project $r_F(1-\tau_K)$, then it is optimal for the entrepreneur to set $k_E = 0$ and to allocate all her capital to her risk-free project, or sell it and use the revenue to lend to a bank. If the after-tax excess return from investing the risky project is positive, i.e. $(\theta r_E - r_F)(1-\tau_K) > 0$, then the entrepreneur allocates an amount of capital to her risky project which is proportional to her initial lifetime resources P. As in many other models with financial market frictions, it therefore follows that the allocation of capital in the economy depends on the wealth distribution across entrepreneurs – capital is not necessarily allocated to its most productive uses. If the after-tax excess return from investing the risky project is sufficiently high, the entrepreneur will choose the largest k_E that guarantees non-negative consumption under the lowest possible realization of ϵ.

All else equal, a richer entrepreneur invests more in her risky project. Furthermore, entrepreneurs invest more in risky projects and less in risk-free projects when: *(i)* the after-tax return to risky projects is relatively higher, *(ii)* the after-tax return to risk-free projects is relatively lower or *(iii)* the agency friction is less severe (i.e. lower ϕ).

Occupational Choice As discussed above, newborn households will be indifferent between becoming entrepreneurs and workers in equilibrium. The lifetime resources of a newborn worker are equal to F^N and the lifetime resources of a newborn entrepreneur are equal to F. Therefore, the occupational indifference condition can be written as:

$$V^N(F^N, X) = \mathbb{E}_\theta V(F, \theta, X).$$

This condition pins down the value of N in equilibrium: if N is very large, then a high supply of labor reduces the steady state wage, thereby increasing profits. Lower wages reduce F^N and higher profits increase F. N adjusts until the value of a newborn worker equals the value of a newborn entrepreneur.

3.4 Aggregate Steady State

Having characterized worker and entrepreneur's choices, we next formally characterize a steady state of the model.

3.4.1 Post-tax Prices

First, we note that in any equilibrium, entrepreneurs' optimal production of final goods implies that:

$$
\begin{aligned}
r_E &= f_1\left(\frac{K_E}{1-N}, \frac{K-K_E}{1-N}, \frac{N}{1-N}\right), \\
r_F &= f_2\left(\frac{K_E}{1-N}, \frac{K-K_E}{1-N}, \frac{N}{1-N}\right), \\
w &= f_3\left(\frac{K_E}{1-N}, \frac{K-K_E}{1-N}, \frac{N}{1-N}\right), \\
\pi_F &= \frac{Y - r_E K_E - r_F(K-K_E) - wN}{1-N},
\end{aligned}
$$

where $f_i\left(\frac{K_E}{1-N}, \frac{K-K_E}{1-N}, \frac{N}{1-N}\right)$ denotes the derivative of $f\left(\frac{K_E}{1-N}, \frac{K-K_E}{1-N}, \frac{N}{1-N}\right)$ with respect to its i^{th} argument.

Since the planner has four tax instruments, τ_K, τ_W, τ_N and τ_C he can set these four tax instruments to target the values of four post-tax prices (subject to its budget constraint). As a consequence of this, it will be useful in the discussion of optimal taxation below to characterize the entire steady state of the model in terms of only resource allocations and four post-tax prices. This can be done using the following definitions of post-tax prices:

$$
\tilde{w} = \frac{w(1-\tau_N)}{1+\tau_C} \tag{17}
$$

$$
\tilde{R}_F = R_F - 1 \tag{18}
$$

$$
\tilde{\pi}_F = \frac{(1-\tau_K)\pi_F}{1+\tau_C}, \tag{19}
$$

$$
\tilde{r}_X = \frac{(r_E - r_F)(1-\tau_K)}{1+\mathcal{R}(K_E, K_F, N, \pi_0)} \tag{20}
$$

where $\mathcal{R}(\cdot)$ denotes the function $\mathcal{R}: \mathbb{R}^4_{++} \to \mathbb{R} \cup \{\infty\}$, which satisfies:

$$
\mathcal{R}(K_E, K, N, \pi_0) = -1 + \frac{(1-N)\pi_0 + K_E}{\left(\frac{(1-N)f(\cdot) - f_1(\cdot)K_E - f_2(\cdot)(K-K_E) - f_3(\cdot)N}{f_1(\cdot) - f_2(\cdot)}\right) + K_E}, \tag{21}
$$

where, abusing notation, $f(\cdot)$ denotes $f\left(\frac{K_E}{1-N}, \frac{K-K_E}{1-N}, \frac{N}{1-N}\right)$ and π_0 is an arbitrary constant. Throughout, we will use $\mathcal{R}$ as a shorthand to denote $\mathcal{R}(K_E, K, N, \pi_0)$. Below, when discussing comparative statics, we typically set π_0 such that $\mathcal{R} = 0$ in some initial steady state.

Intuitively, $\tilde{w}$, $\tilde{R}_F$, $\tilde{\pi}_F$ and $(1+\mathcal{R})\tilde{r}_X$ represent, respectively, the post-tax wage rate, the post tax risk-free rate, the entrepreneur's post-tax profit from selling final goods, and the post-tax excess return to high productivity entrepreneurs from using the risky technology. It is useful to define the latter as $(1+\mathcal{R})\tilde{r}_X$ rather than simply $\tilde{r}_X$ because theplanner's ability to control all post-tax returns is limited due to only having four tax instruments. As such, it will be convenient to think of $\tilde{r}_X$ intuitively as the part of the excess return from the risky technology that the planner can choose by varying taxes, and $1+\mathcal{R}$ as the part of the excess return to risky capital that the planner does not choose, and instead varies endogenously due to variation in pre-tax prices. Pre-tax prices in turn depend on K_E, K and N, and so $\mathcal{R}$ depends on K_E, K and N. At $\mathcal{R}=0$, as we typically assume, it can readily be shown that $\mathcal{R}$ is decreasing in K_E, holding constant K and N, which corresponds to the intuition that a higher level of capital in the risky technology reduces the excess return to this technology.

To interpret $\mathcal{R}$, we combine (20) with the definition of π_F above and use (19) to obtain

$$\tilde{\pi}_F = \frac{\tilde{r}_X}{1+\tau_C}\left(\pi_0 - \frac{\mathcal{R}K_E}{1-N}\right).$$

The aggregate post-tax profits of entrepreneurs can therefore be written:

$$(1+\tau_C)\tilde{\pi}_F(1-N) + \tilde{r}_X(1+\mathcal{R})K_E + \tilde{R}_F(K-A^N) = (\pi_0(1-N)+K_E)\tilde{r}_X + \tilde{R}_F(K-A^N)$$

where A^N is the aggregate wealth of workers and $K-A^N$ is the aggregate wealth of entrepreneurs.

This equation reveals that an increase in $\mathcal{R}$ does not, by itself directly affect entrepreneurs profits at all. Rather, it raises the excess return to capital, $\tilde{r}_X(1+\mathcal{R})$ and reduces entrepreneurs' profits from selling final goods, $\tilde{\pi}_F$. Thus, $\mathcal{R}$ represents the degree to which entrepreneurs' profts reflect risky investments, rather than simply the rents $\tilde{\pi}_F$ which all entrepreneurs earn equally from selling final goods. As such, an increase in $\mathcal{R}$ does not directly affect the income of entrepreneurs in total, but it increases the income of wealthy high ability entrepreneurs (who choose high levels of k_E) at the expense of low ability entrepreneurs. Since $\mathcal{R}$ does not reflect total entrepreneurial profits, then, but simply the allocation of entrepreneurial profits, it is not directly affected by the planner's tax instruments, but changes endogenously as a consequence of changes in pre-tax prices when the levels of K_E, K and N change.

3.4.2 Steady State Characterization

Having defined post-tax prices, we now formally characterize the steady state of the model. We define a steady state as follows:

Definition 2. *A* ***steady state*** $\mathscr{S}$ *of the economy is a set of values of tax rates* $\{\tau_W^*, \tau_K^*, \tau_C^*, \tau_N^*\}$, *prices* $\{r_E^*, r_F^*, w^*\}$, *aggregate variables* $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}\right\}, \mathbb{P}$ *and an equilibrium* $\mathcal{E}$

in which all tax rates, prices and aggregate variables are equal to these steady state values in every period.

We focus on steady states in which no capital is hidden because, as argued below, the planner will never design a tax policy to select a steady state in which this happens. The full set of conditions that must be satisfied in a steady state are summarized in Proposition 2 below.

Proposition 2. *There exists a steady state $\mathscr{S}$ which is consistent with the particular values of aggregate variables $\{K^*, K_E^*, C^*, N^*, \mathcal{R}^*, Y^*\}$, post-tax prices $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ and consumption tax rate τ_C^* and in which no entrepreneurs hide capital or intermediate goods, if and only if the following conditions hold:*

$$C^* = \frac{\rho+\gamma}{\tilde{R}_F^*+\gamma}\left(Y^* - \delta K^* - \overline{G} - \frac{\tilde{r}_X^*(1+\mathcal{R}^*)K_E^*}{1+\tau_C^*} + \frac{\gamma K^*}{1+\tau_C}\right) \tag{22}$$

$$Y^* - \delta K^* - \overline{G} = C^* = N^*\tilde{w}^* + (1-N^*)\tilde{\pi}_F^* + \frac{\tilde{r}_X^*(1+\mathcal{R}^*)K_E^*}{1+\tau_C^*} + \frac{\tilde{R}_F^* K^*}{1+\tau_C^*} \tag{23}$$

$$\log(\tilde{w}^*) = \log(\tilde{\pi}_F^*) + \frac{g(1)}{\rho+\gamma}\left[(1+\mathcal{R}^*)\,\tilde{r}_X\hat{k}(1) - \frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(1)\varphi\right)^2\right] \tag{24}$$

$$K_E = \frac{\tilde{r}_X(1+\mathcal{R}^*)K_E + \frac{\gamma(1+\tau_C)}{\gamma+\tilde{R}_F}\left(\tilde{\pi}_F^*(1-N^*)\right)}{\rho+\gamma-\tilde{R}_F}\hat{k}_E(1)\mu(1) \tag{25}$$

$$\mu(1) = \frac{\lambda_\theta g(1) + \frac{\gamma g(1)(1+\tau_C)}{\gamma+\tilde{R}_F}\left(\tilde{\pi}_F^*(1-N^*)\right)\frac{\hat{k}_E(1)\mu(1)}{K_E}}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^* - (1+\mathcal{R}^*)\tilde{r}_X^*\hat{k}_E(1)}, \tag{26}$$

$$\tilde{\pi}_F^* = \left(\frac{\tilde{r}_X^*}{1+\tau_C^*}\right)\left(\pi_0 - \frac{\mathcal{R}^* K_E^*}{1-N^*}\right) \tag{27}$$

where $\mathcal{R}^* = \mathcal{R}(K_E^*, K^*, N^*, \pi_0)$, $Y^* = f\left(\frac{K_E^*}{1-N^*}, \frac{K^*-K_E^*}{1-N^*}, \frac{N^*}{1-N^*}\right)(1-N^*)$ and where $K_E^* < K^*$, $\frac{1}{f_1^*-f_2^*} > \frac{\phi(1+\tau_C^*)}{\tilde{r}_X^*} > 0$ and $\hat{k}_E(\theta)$ is given by equation (14).

Proof. See Appendix B.4. □

Note that none of the equations in Proposition 2 make any reference to pre-tax prices or to tax rates, apart from τ_C; they make reference only to post-tax prices, τ_C and $\mathcal{R}$. Therefore, as long as the inequality conditions in the proposition are satisfied, changes in pre-tax prices have no effect on steady state allocations, conditional on the value of post-tax prices and $\mathcal{R}$. $\mathcal{R}$ matters for the steady state allocation because, as discussed above, a higher level of

$\mathcal{R}$ entails a redistribution of profits from low ability entrepreneurs, to entrepreneurs that choose high levels of capital in the risky technology.

The interpretation of the equilibrium conditions in Proposition 2 is intuitive. The first equilibrium condition is the consumption function for the economy. Each agents' marginal propensity to consume out of earnings is $MPC \equiv \frac{\rho+\gamma}{\tilde{R}_F^*+\gamma}$ in the steady state, which depends on preferences and the risk-free rate. Agents consume based on their net resources after tax and depreciation, $Y^* - \delta K^* - \overline{G}$. Additionally, entrepreneurs do not consume out of their risky capital income, since a higher risky capital income means a higher return to saving. Log utility implies that all this extra capital income is saved. Finally, agents consume out of the after-tax resources they earn from annuities each period, $\frac{\gamma K}{1+\tau_C}$. Equation (23) states that output net of investment and government spending equals consumption, which also must equal the sum of post-tax resources of all agents in the economy, since saving (net of depreciation) must equal zero in the steady state. Equation (24) is the condition under which agents are indifferent about their occupation in the steady state and equations (26) and (25) arise from the stationary wealth distribution that is induced by entrepreneurs' policy choices. Lastly, equation (27) was derived in the previous section.

3.5 The Effect of Taxes on Equilibrium Allocations

In Section 4.2 we will show that optimal steady state tax rates can be written in terms of estimatable "sufficient statistics". These sufficient statistics are elasticities of aggregate variables with respect to tax changes. In doing that, our approach provides a bridge between theory and empirics in the spirit of Chetty (2008), Piketty and Saez (2013), Piketty, Saez and Stantcheva (2014) and Saez and Stantcheva (2018), where the sufficient statistics are behavioral elasticities that capture the response of aggregate outcomes to small changes in taxes. To that end, in this section, we formally define the elasticities that shape the optimal taxes and use them to gain intuition about the effect of the various tax instruments on key aggregate variables in the economy.

As changes in taxes in this economy have a wide range of general equilibrium effects, to maximize the clarity of exposition, we proceed in two steps. In the next two sections, we characterize the partial equilibrium elasticity with respect to tax changes of aggregate output Y, total capital stock K and the capital invested in the entrepreneurial sector K_E, holding pre-tax prices constant. In Section 3.5.2 we consider the general equilibrium effects of tax changes.

3.5.1 Partial Equilibrium Effects of Taxes on Equilibrium Allocations

To study the partial equilibrium effect of changes in tax rates on aggregate output Y, total capital stock K and the capital invested in the entrepreneurial sector K_E, we consider the long run effect on the steady state of a small change in tax rates holding constant pre-tax prices and the fraction N of agents who choose to become workers. By holding pre-tax prices completely constant, we will hold $\mathcal{R}(K, K_E, N, \pi_0)$ constant, since the latter captures the effects of varying pre-tax prices on profits and allocations. In the next section, we will discuss the additional general equilibrium effects that arise when N and $\mathcal{R}(K, K_E, N, \pi_0)$ can vary as the values of K, K_E and N change in response to changes in taxes. To emphasize the effect that financial frictions have on the determination of optimal taxes, we assume throughout that the planner chooses the levels of τ_K, τ_W and adjusts the labor income tax τ_N to ensure that its budget balances in the steady state, while τ_C remains unchanged. For ease of interpretation, we fix $\mathcal{R} = 0$ by setting the arbitrary constant π_0 so that $\mathcal{R}^* = \mathcal{R}(K^*, K_E^*, N^*, \pi_0) = 0$ in the initial steady state, before any tax changes take effect.[12]

We formally define long-run elasticities as follows.

Definition 3. *The partial equilibrium elasticities of Y, K and K_E with respect to the tax rate τ_K are defined as:*

$$e^Y_{\tau_K} \equiv \frac{(1-\tau_K)}{Y}\frac{\partial Y}{\partial \tau_K}\Big|_{\mathcal{R}=0,N} \qquad e^K_{\tau_K} \equiv \frac{(1-\tau_K)}{K}\frac{\partial K}{\partial \tau_K}\Big|_{\mathcal{R}=0,N} \qquad e^{K_E}_{\tau_K} \equiv \frac{(1-\tau_K)}{K_E}\frac{\partial K_E}{\partial \tau_K}\Big|_{\mathcal{R}=0,N}.$$

The partial equilibrium (semi-)elasticities of Y, K and K_E with respect to the tax rate τ_W are defined as:

$$e^Y_{\tau_W} \equiv \frac{1}{Y}\frac{\partial Y}{\partial \tau_W}\Big|_{\mathcal{R}=0,N} \qquad e^K_{\tau_W} \equiv \frac{1}{K}\frac{\partial K}{\partial \tau_W}\Big|_{\mathcal{R}=0,N} \qquad e^{K_E}_{\tau_W} \equiv \frac{1}{K_E}\frac{\partial K_E}{\partial \tau_W}\Big|_{\mathcal{R}=0,N}.$$

That is, $e^Y_{\tau_K}$ is the elasticity of aggregate steady state output with respect to the tax rate τ_K, holding constant pre-tax prices (including the value of $\mathcal{R}(K, K_E, N)$) and the fraction N of agents who choose to become workers, and assuming that τ_N adjusts to balance the government's budget, and $e^Y_{\tau_W}$ is the corresponding (semi-)elasticities of aggregate output with respect to τ_W. The elasticities of aggregate and risky capital are defined in a similar fashion.

By considering small perturbations in tax rates around the conditions that characterize a

[12] Using the definition of $\mathcal{R}$ above and noting the each factor is paid its marginal product in equilibrium, it follows that this will be true if: $\pi_0 = \frac{\pi_F^*}{r_E^* - r_F^*}$. Holding $\mathcal{R}$ fixed at zero makes intepretation easier, since it means that we can interpret changes in $\tilde{r}_X$ as corresponding to changes in τ_K, holding pre-tax prices constant. $\mathcal{R} = 0$ is necessary for this interpretation, since $\mathcal{R} = 0$ ensures that, $\tilde{\pi}_F = \frac{\tilde{r}_X \pi_0}{1+\tau_C}$. Then, a decrease in $\tilde{r}_X$ leads to a proportional decrease in $\tilde{\pi}_F$, which is the same effect as an increase in τ_K holding pre-tax prices constant.

steady state of the economy, all the elasticities defined above can be characterized in closed form, as functions of structural parameters and aggregate variables. We relegate the details of this derivation to Appendix B.5 and directly provide the characterization of the elasticities of Y, K_E and K with respect to taxes in Proposition 3 below.

Proposition 3. *The elasticity of aggregate output Y with respect to the tax rates τ_K and τ_W, holding constant pre-tax prices and the fraction N of agents who choose to become workers, and assuming that τ_N adjusts to balance the government's budget is, respectively:*

$$\begin{aligned} e^{Y}_{\tau_K} &= (r_E - r_F)\frac{K_E}{Y}e^{K_E}_{\tau_K} + r_F\frac{K}{Y}e^{K}_{\tau_K} \\ e^{Y}_{\tau_W} &= (r_E - r_F)\frac{K_E}{Y}e^{K}_{\tau_W} + r_F\frac{K}{Y}e^{K}_{\tau_W}. \end{aligned}$$

Proof. The result follows from partially differentiating equation (23) with respect to $\tilde{r}_X$, $\tilde{R}_F$ and τ_C, combining this with Definition 3 and rearranging. □

The expression for $e^{Y}_{\tau_K}$ reveals that a change in tax rates affects aggregate output via its effect on aggregate capital accumulation ($e^{K}_{\tau_K}$) and also by affecting how far capital is allocated to the risky technology ($e^{K_E}_{\tau_K}$). An increase in the fraction of capital allocated to the risky technology increases aggregate output, because $r_E - r_F > 0$ must hold in equilibrium, since entrepreneurs must receive a higher return to the risky technology to compensate for risk. A decrease in capital income taxes (with a compensating rise in τ_N) will typically raise aggregate steady state output, holding constant N and pre-tax prices. This is because a decrease in capital income taxes tends to increase aggregate capital accumulation (so $e^{K}_{\tau_K} < 0$) and encourages a larger quantity of capital in the risky technology (so $e^{K_E}_{\tau_K} < 0$). We verify that $e^{K}_{\tau_K} < 0$ and $e^{K_E}_{\tau_K} < 0$ in Propositions 4 and 5, which characterize the partial equilibrium effects of tax changes on K_E and K, respectively.

Proposition 4. *The elasticity of the risky capital stock K_E with respect to the tax rates τ_K and τ_W, holding constant pre-tax prices and the fraction N of agents who choose to become workers, and assuming that τ_N adjusts to balance the government's budget is, respectively:*

$$\begin{aligned} e^{K_E}_{\tau_K} &= (e^{\hat{k}_E}_{\tau_K} + e^{\mu}_{\tau_K} - 1)M_{K_E} - (1-\tau_K)(r_F - \delta)M_{\tilde{R}_F} \\ e^{K_E}_{\tau_W} &= (e^{\hat{k}_E}_{\tau_W} + e^{\mu}_{\tau_W})M_{K_E} - M_{\tilde{R}_F}, \end{aligned}$$

where

$$\begin{aligned} M_{K_E} &= 1 + \frac{K_E \tilde{r}_X(\gamma + \tilde{R}_F)}{(1+\tau_C)(1-N)\tilde{\pi}_F \gamma} \\ M_{\tilde{R}_F} &= \frac{-1}{\gamma + \tilde{R}_F} + \frac{M_{K_E}}{\rho + \gamma - \tilde{R}_F} \\ e_{\tau_K}^{\hat{k}_E} &= \frac{1-\tau_K}{\hat{k}_E(1)} \frac{\partial \hat{k}_E(1)}{\partial \tau_K} \\ e_{\tau_j}^{\mu} &= \frac{1-\tau_K}{\mu(1)} \frac{\partial \mu(1)}{\partial \tau_K}, \end{aligned}$$

with analogous definitions of $e_{\tau_W}^{\hat{k}_E}$ *and* $e_{\tau_W}^{\mu}$*. Moreover,* $e_{\tau_K}^{K_E} < 0$*.*

Proof. The expressions for $e_{\tau_K}^{K_E}$ and $e_{\tau_W}^{K_E}$ follow from partially differentiating equation (25) with respect to $\tilde{r}_X$ and $\tilde{R}_F$, combining this with Definition 3 and rearranging. To show that $e_{\tau_K}^{K_E} < 0$ suppose otherwise, that K_E is (weakly) increasing in τ_K. Note that $e_{\tau_K}^{\hat{k}_E} < 0$ from the characteriation of $\hat{k}_E$ in (14). Furthermore, if K_E is increasing in τ_K, the equation (26) for $\mu(1)$ implies that $\mu(1)$ is decreasing in τ_K so that $e_{\tau_K}^{\mu} < 0$. Then, since $-1 - \frac{\tilde{R}_F}{\gamma + \tilde{R}_F} < 0$, it follows that $e_{\tau_K}^{K_E} < 0$, which is the required contradiction. □

Proposition 5. *The elasticity of the aggregate capital stock K with respect to the tax rates τ_K and τ_W, holding constant pre-tax prices and the fraction N of agents who choose to become workers, and assuming that τ_N adjusts to balance the government's budget is, respectively:*

$$e_{\tau_j}^{K} = \frac{\left(\frac{K_E}{K}\right) e_{\tau_j}^{K_E} \left[(r_E - r_F)(1 - MPC) + \frac{\tilde{r}_X}{(1+\tau_C)} MPC \right] + e_{\tau_j}^{S_{SUB}}}{-(r_F - \delta)(1 - MPC) + \left(\frac{\gamma}{1+\tau_C}\right) MPC},$$

for $j \in \{K; W\}$, where

$$\begin{aligned} e_{\tau_K}^{S_{SUB}} &= -\left(\frac{\tilde{r}_X K_E}{K(1+\tau_C)}\right) MPC - \frac{C}{K} e_{\tau_K}^{MPC}, \\ e_{\tau_W}^{S_{SUB}} &= -\frac{C}{K} e_{\tau_W}^{MPC}, \end{aligned}$$

and $MPC = \frac{\rho+\gamma}{\tilde{R}_F^* + \gamma}$ is the marginal propensity to consume out of aggregate resources and where

$$e_{\tau_j}^{MPC} = \left(\frac{1-\tau_K}{MPC}\right) \frac{\partial MPC}{\partial \tau_j}.$$

Furthermore, $e_{\tau_K}^{MPC} > 0$. If the MPC is sufficiently close to 1, then $e_{\tau_K}^{K} < 0$.

Proof. The expression for $e_{\tau_j}^{K}$ follows from partially differentiating the steady state aggregate resource constraint (23) with respect to $\tilde{r}_X$ and $\tilde{R}_F$, combining this with Definition 3, using the consumption function (22) to substitute for the changes in aggregate consumption and rearranging. Taking the derivative $\frac{\partial MPC}{\partial \tau_K}$ immediately reveals that $e^{MPC}_{\tau_K} = \frac{r_F}{\tilde{R}_F+\gamma} > 0$. To show that an MPC close enough to 1 implies that $e^{MPC}_{\tau_K} < 0$, note that $e^{K_E}_{\tau_K} < 0$ from Proposition 4. Additionally, $\tilde{r}_X > 0$ must hold in the steady state or no entrepreneur would put capital into the risky technology. Then, for an MPC close enough to 1, the numerator in the expression for $e^{K}_{\tau_j}$ must be negative. Equally, for an MPC close enough to 1, the denominator of the expression for $e^{K}_{\tau_j}$ must be positive. □

Above we established that the effect of a capital income tax on aggregate output depends Y on its effect on K and K_E. The results in Propositions 4 and 5 allow us to inspect this effect in more detail. Focusing first on Proposition 4, we see that a tax change has four effects on aggregate K_E. First, it affects $\hat{k}_E$, that is, the fraction of each entrepreneur's resources that they choose to put in the risky technology. Second, it affects $\mu(1)$ – the fraction of resources held by high ability entrepreneurs, which affects K_E since low ability entrepreneurs do not use the risky technology. These first two effects are captured by the terms $e^{\hat{k}_E}_{\tau_j} + e^{\mu}_{\tau_j}$ for each $j \in \{K; W\}$. Third, a change in capital tax rate τ_K directly reduces the post-tax profits of entrepreneurs, which reduces the resources they have to allocate to the risky technology – this effect is captured by the "-1" term in the expression for $e^{K_E}_{\tau_K}$. These first three effects have an impact on K_E which is multiplied by the term M_{K_E}. $M_{K_E} > 1$ represents the effect that, as entrepreneurs devote more resources to the risky technology, this in turn increases their incomes, allowing them to devote even larger amounts of resources to the risky technology. The fourth effect is that changes in capital income and wealth taxes affect the total post-tax return $\tilde{R}_F$ to the risk-free technology. This effect is captured by the terms mutliplying $M_{\tilde{R}_F}$. $M_{\tilde{R}_F}$ in turn has two components: a term $\frac{-1}{\gamma+\tilde{R}_F}$ capturing the fact that a fall in $\tilde{R}_F$ causes a substitution effect hat increases K_E, and a term multiplying M_{K_E} which captures the effect that a change in $\tilde{R}_F$ changes entrepreneurs' total resources, which in turn affects the aggregate level of K_E.

Now, considering Proposition 5, the numerator in the expression for $e^{K}_{\tau_j}$ shows the degree to which an increase in taxes raises aggregate saving, holding constant K, and the denominator shows how rapidly aggregate saving decreases as K rises. The terms multiplying $e^{K_E}_{\tau_j}$ in the numerator show that a major effect of a change in taxes on aggregate saving is the effect of the tax change on K_E. An increase in K_E tends to increase aggregate savings through two channels. First, it increases output by an amount in proportion to $(r_E - r_F)$, which increase saving in proportion to $(r_E - r_F)(1 - MPC)$. Second, when entrepreneurs hold higher higher levels of K_E, this implies their marginal return to saving is higher, which further increases their saving in proportion to $\frac{r_E(1-\tau_K)}{(1+\tau_C)} MPC$. Finally, a change in taxes also affects saving via substitution effects, since taxes directly affect the post-tax return to saving. The size of this substiution effect is given by $e^{S_{SUB}}_{\tau_j}$. For realistic parameter values, the MPC is relatively close to 1 (since this is the marginal propensity to consume out of steady state income), which implies that the denominator is small and positive, indicating that a change in the aggregate capital stock has a relatively small effect on aggregate saving. Therefore, increases in capital income taxes have the potential to cause substantial decreases in steady state aggregate capital K.

3.5.2 General Equilibrium Effects of Taxes on Equilibrium Allocations

Tax changes in the model have a wide range of general equilibrium effects on aggregate variables, which are naturally more complicated to analyze than the partial equilibrium effects. Fortunately, most of these general equilibrium effects can be ignored when finding optimal tax rates. Recall that the equations of Proposition 2 imply that, conditional on post-tax prices, pre-tax prices only affect steady state allocations via $\mathcal{R}$, with $\mathcal{R}$ affecting the return to capital and profits. Then, since the planner can set tax rates to target particular values of post-tax prices (subject to its budget constraint), it emerges that the only way changes in pre-tax prices matter for the optimal allocation and, therefore, optimal tax rates, is via the function $\mathcal{R}$. That is, optimal taxes can be correctly calculated *as if* tax changes had no effect on pre-tax prices except via the effect of $\mathcal{R}$ on the return to capital and profits. This is shown in Section 4.2.[13] Since $\mathcal{R}$ reflects the distribution of profits between different entrepreneurs, it emerges that the only way changes in pre-tax prices matter for optimal taxation is via their effect on the distribution of profits across entrepreneurs. As such, in this section, we study the effects of tax changes when tax changes also affect N, as well as the pre-tax return to capital and profits via $\mathcal{R}$, but we ignore all other effects of tax changes on pre-tax prices. Thus, we let $\overline{e}^{Y}_{\tau_K}$ denote the elasticity of Y with respect to τ_K when $\mathcal{R}$ and N may change, but pre-tax prices are otherwise held fixed. As a shorthand, we call this, the general equilibrium elasticity of Y with respect to τ_K. Analogously with Definition 3, we define $\overline{e}^{Y}_{\tau_K}$ according to:

$$\overline{e}^{Y}_{\tau_K} = \frac{1-\tau_K}{Y}\frac{\partial Y}{\partial \tau_K}.$$

We may then characterize $\overline{e}^{Y}_{\tau_K}$ following the same procedure used to characterize $e^{Y}_{\tau_K}$. We obtain

$$\overline{e}^{Y}_{\tau_K} = (r_E - r_F)\frac{K_E}{Y}\overline{e}^{K_E}_{\tau_K} + r_F\frac{K}{Y}\overline{e}^{K}_{\tau_K} + (w-\pi_F)\overline{e}^{N}_{\tau_K},$$

where $\overline{e}^{K}_{\tau_K}$, $\overline{e}^{K_E}_{\tau_K}$ and $\overline{e}^{N}_{\tau_K}$ are defined analogously to the partial equilibrium elasticities; that is, $\overline{e}^{K}_{\tau_K} = \frac{1-\tau_K}{K}\frac{\partial K}{\partial \tau_K}$, $\overline{e}^{K_E}_{\tau_K} = \frac{1-\tau_K}{K_E}\frac{\partial K_E}{\partial \tau_K}$, and $\overline{e}^{N}_{\tau_K} = \frac{1-\tau_K}{N}\frac{\partial N}{\partial \tau_K}$.

This equation indicates that the effect of a small change in τ_K on aggregate output is ambiguous once changes in N are allowed for: a small cut in τ_K with a corresponding rise in

[13] Similar findings that general equilibrium effects can be ignored in optimal tax formulae also hold and are discussed in earlier related work in the litererature, such as Piketty and Saez (2013).

τ_N will tend to increase capital accumulation, but will also tend to discourage agents from becoming workers, which could reduce output if $w > \pi_F$.

Defining the general equilibrium elasticity with respect to τ_W in the obvious analogous way, and repeating the same steps that were used to characterize $\bar{e}^Y_{\tau_K}$, we obtain:

$$\bar{e}^Y_{\tau_W} \quad = \quad (r_E - r_F)\frac{K_E}{Y}\bar{e}^{K_E}_{\tau_W} + r_F\frac{K}{Y}\bar{e}^K_{\tau_W} + (w - \pi_F)\bar{e}^N_{\tau_W}.$$

Repeating the same approach to characterize the general equilibrium elasticities $\bar{e}^K_{\tau_j}$, and $\bar{e}^{K_E}_{\tau_j}$, for for each $j \in \{K; W\}$, reveals that these are also very similar to the partial equilibrium elasticities, but with additional terms reflecting the change in $\mathcal{R}$ and change in N after a change in taxes. Specifically, for each $j \in \{K; W\}$

$$\bar{e}^{K_E}_{\tau_j} = e^{K_E}_{\tau_j} - \frac{N}{1-N}\bar{e}^N_{\tau_K} + \left(\frac{(M_{K_E} - 1)\tilde{R}_F}{\gamma + \tilde{R}_F}\right)(\mathcal{R}'_1(\cdot)\bar{e}^{K_E}_{\tau_j} + \mathcal{R}'_2(\cdot)\bar{e}^K_{\tau_j} + \mathcal{R}'_3(\cdot)\bar{e}^N_{\tau_j})$$

and

$$\begin{aligned}\bar{e}^K_{\tau_j} \quad &= \quad e^K_{\tau_j} + \frac{\left(\frac{K_E}{K}\right)\left(\bar{e}^{K_E}_{\tau_j} - e^{K_E}_{\tau_j}\right)\left[(r_E - r_F)(1 - MPC) + \frac{\tilde{r}_X}{(1+\tau_C)}MPC\right]}{-(r_F - \delta)(1 - MPC) + \left(\frac{\gamma}{1+\tau_C}\right)MPC} \\ &+ \quad \frac{[(w - \pi)(1 - MPC)]\frac{N}{K}\bar{e}^N_{\tau_j} + \left[\frac{K_E}{1+\tau_C}\right]\frac{\tilde{r}_X}{K}MPC(\mathcal{R}'_1(\cdot)\bar{e}^{K_E}_{\tau_j} + \mathcal{R}'_2(\cdot)\bar{e}^K_{\tau_j} + \mathcal{R}'_3(\cdot)\bar{e}^N_{\tau_j})}{-(r_F - \delta)(1 - MPC) + \left(\frac{\gamma}{1+\tau_C}\right)MPC},\end{aligned}$$

where M_{K_E} is as defined in Proposition 4, and $\mathcal{R}'_i(\cdot)$ denotes the derivative of $\mathcal{R}(K_E, K, N, \pi_0)$ with respect to its ith argument, where π_0 is set such that $\mathcal{R} = 0$ in the initial steady state.

The equation for $\bar{e}^{K_E}_{\tau_j}$ shows that tax changes affect K_E through two additional general equilibrium effects, in addition to the partial equilibrium effects noted above. First, a tax change affects the number of agents who choose to become workers, which affects K_E becomes workers do not use the risky technology. Second, a tax change affects pre-tax prices via $\mathcal{R}$. The change in $\mathcal{R}$ after the tax change is equal to $\mathcal{R}'_1(\cdot)\bar{e}^{K_E}_{\tau_j} + \mathcal{R}'_2(\cdot)\bar{e}^K_{\tau_j} + \mathcal{R}'_3(\cdot)\bar{e}^N_{\tau_j})$. A higher $\mathcal{R}$ encourages capital accumulation because a higher $\mathcal{R}$ raises the excess return to capital in the risky technology (which is $\tilde{r}_X(1 + \mathcal{R})$), providing more resources and incentive for entrepreneurs to accumulate higher levels of K_E. The equation for $\bar{e}^K_{\tau_j}$ shows that this is equal to $e^K_{\tau_j}$ plus three additional terms that reflect, first, that in general equilibrium, K_E changes after tax changes by the additional amount $(\bar{e}^{K_E}_{\tau_j} - e^{K_E}_{\tau_j})$; second, that a tax change affects N in proportion to $\bar{e}^N_{\tau_j}$, which affects aggregate income and therefore saving insofar as the income of workers and entrepreneurs differs, and, third, that tax changes affect pre-tax prices via $\mathcal{R}$, with a higher $\mathcal{R}$ representing an increase in the excess return to capital in the risky technology, which encourages saving.

4 Optimal Taxes

In this section, we formulate and solve the planner's problem of choosing optimal taxes in order to maximize the steady-state welfare of a newborn agent. In solving for the optimal taxes we use a calculus of variations approach, which requires that we first characterize the the effects of changes in tax rate on welfare. It is then straightforward to use the effects of tax changes on welfare to deduce optimal tax rates, since the first order conditions for optimal tax rates are simply that the marginal effects of τ_K and τ_W on welfare are all equal to zero.

4.1 Effects of Tax Changes on Welfare

The measure of welfare we consider is the lifetime present discounted utility of a newborn agent in the steady state, which we denote by $\mathcal{W}$. Since newborn agents are indifferent in equilibrium between becoming entrepreneurs and workers, this is the same as the lifetime present discounted utility of a newborn worker in the steady state. In discrete time, this is

$$\mathcal{W} = \sum_{s=0}^{\infty} (1-\rho)^s (1-\gamma)^s \log\left(c_s^N\right).$$

We can derive a simple formula of the change in welfare that results from a marginal change in tax rates. As is common in the literature, we focus on the consumption equivalent welfare change. Considering a small change in tax rates that changes steady state worker consumption by $\mathrm{d}c_s^N$ in each period s of a newborn worker's life, the consumption equivalent change in welfare Δ^N satisfies:

$$\sum_{s=0}^{\infty} (1-\rho)^s (1-\gamma)^s \log\left(c_s^N (1+\Delta^N)\right) = \sum_{s=0}^{\infty} (1-\rho)^s (1-\gamma)^s u(c_s^N + \mathrm{d}c_s^N).$$

Envelope theorem arguments, which we formalize in Appendix B.6, allow us to obtain a simple formula for Δ^N. In particular, Δ^N satisfies:

$$\tilde{w}\Delta^N = \mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \frac{\mathcal{A}^N}{1+\tau_C},$$

where $d\tilde{w}$ and $d\tilde{R}_F$ are the change in $\tilde{w}$ and $\tilde{R}_F$ as a result of the tax change and

$$\mathcal{A}^N = \frac{\sum_{s=0}^{\infty} \left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s a_s^N}{\sum_{s=0}^{\infty} \left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s}$$

is the average value of the worker's discounted lifetime assets.

It is possible to extend this argument to the continuous time case, as the following lemma shows.

Lemma 4. *The consumption equivalent change in steady state welfare Δ^N from a marginal change in taxes satisfies:*

$$\tilde{w}\Delta^N = d\tilde{w} + \frac{\mathcal{A}^N d\tilde{R}_F}{1+\tau_C}$$

where

$$\mathcal{A}^N = (\gamma + \tilde{R}_F) \int_{s=0}^{\infty} e^{-(\gamma+\tilde{R}_F)s} a_s^N ds.$$

Proof. See Appendix B.6 □

Intuitively, the consumption equivalent change in welfare is simply equal to the change in worker income on the margin due to changes in $\tilde{w}$ and $\tilde{R}_F$, divided by the worker's average lifetime earnings $\tilde{w}$.

Having established a formula for the consumption equivalent change in welfare, we may study the effects of tax changes on welfare using the same approach we took in the previous section.

Partial Equilibrium Effect of Tax Changes on Welfare

To build intuition, we first provide a heuristic derivation of the marginal effect of a change in τ_K on welfare in partial equilibrium. We assume that the planner changes τ_K by some amount $d\tau_K$, and changes τ_N by some amount $d\tau_N$ to balance the budget. Using the definitions of post-tax prices in Section 3.4.1 (and keeping pre-tax prices fixed for a partial equilibrium analysis) and using Lemma 4, the resulting consumption equivalent change in welfare satisfies:

$$\Delta^N = -\frac{d\tau_N}{1-\tau_N} - \left(\frac{(r_F - \delta)\mathcal{A}^N}{(1-\tau_N)w}\right) d\tau_K.$$

We use the government's budget constraint to infer $d\tau_N$ as a function of $d\tau_K$. To that end, it is convenient to write the government's budget constraint as

$$\overline{G} = \tau_N B_{\tau_N} + \tau_K B_{\tau_K} + \tau_W B_{\tau_W} + \tau_C B_{\tau_C},$$

where B_{τ_j} is the tax base for the tax τ_j, so that $B_{\tau_N} = wN$, $B_{\tau_K} = (1-N)\pi_F + (r_E - r_F)K_E + (r_F - \delta)K$, $B_{\tau_W} = K$ and $B_{\tau_C} = C$.

Differentiating the government's budget constraint with respect to $\mathrm{d}\tau_K$, holding constant pre-tax prices, $\mathcal{R}$ and N (for a partial equilibrium analysis) we obtain that

$$0 = B_{\tau_N}\mathrm{d}\tau_N + B_{\tau_K}\mathrm{d}\tau_K + \sum_j \tau_j \frac{\partial B_{\tau_j}}{\partial \tau_K}\Big|_{\mathcal{R}=0,N}\mathrm{d}\tau_K.$$

Substituting this into our expression for Δ^N above, multiplying both sides by $\frac{(1-\tau_N)wN}{\mathrm{d}\tau_K}$ and rearranging, we obtain:

$$(1-\tau_N)wN\frac{\partial \Delta^N}{\partial \tau_K}\Big|_{\mathcal{R}=0,N} = B_{\tau_K} + \sum_{j\in\{K;W;C;N\}} \tau_j \frac{\partial B_{\tau_j}}{\partial \tau_K}\Big|_{\mathcal{R}=0,N} - B^N_{\tau_K}N,$$

where $B^N_{\tau_K} = (r_F - \delta)\mathcal{A}^N$ denotes the lifetime average additional tax payments an individual worker would have to make, all else equal, after a unit rise in τ_K.

Intuitively, this equation states that the change in a worker's welfare from an increase in τ_K is proportional to revenue gained from the rise in τ_K (since the revenue gain funds decreases in τ_N) minus the component of the τ_K tax rise that is paid for on average by workers over their lifetime. The revenue gain per unit rise in τ_K is the tax base B_{τ_K}, plus a (negative) term representing the loss in revenue that arises from the behavioural changes caused by the rise in τ_K.

Following the same approach for changes in τ_W, we obtain a formula for the partial equilibrium effect of any tax change on welfare.

Proposition 6. *The partial equilibrium effect of a change in tax rate τ_j on welfare, for $j \in \{K;W\}$ satisfies:*

$$(1-\tau_N)wN\frac{\partial \Delta^N}{\partial \tau_j}\Big|_{\mathcal{R}=0,N} = B_{\tau_j} + \sum_{m\in\{K;W;C;N\}} \tau_m \frac{\partial B_{\tau_m}}{\partial \tau_j}\Big|_{\mathcal{R}=0,N} - B^N_{\tau_j}N.$$

where $B^N_{\tau_K} = (r_F - \delta)\mathcal{A}^N$, $B^N_{\tau_W} = \mathcal{A}$ and $B^N_{\tau_C} = w(1-\tau_N)$.

Proof. Differentiate the aggregate resource constraint in Proposition 2 with respect to $\tilde{R}_F$ and $\tilde{r}_X$, fixing $\mathcal{R} = 0$ and holding constant N. Substitute in the definitions of B_j above, the equation for Δ^N in Lemma 4 and rearrange. □

General Equilibrium Effect of Tax Changes on Welfare

As before, we consider the general equilibrium change allowing $\mathcal{R}$ and N to vary, but otherwise assuming pre-tax prices are fixed. It is straightforward to extend the results of Proposition 6 to the general equilibrium case: the general equilibrium welfare change has the same formula as the partial equilibrium effects, except with general equilibrium derivatives replacing partial equilibrium ones. That is, the general equilibrium effect of a change in tax rate τ_j on welfare, for $j \in \{K; W\}$ satisfies:

$$(1-\tau_N)wN\frac{\partial \Delta^N}{\partial \tau_j} = B_{\tau_j} + \sum_{m\in\{K;W;C;N\}} \tau_m \frac{\partial B_{\tau_m}}{\partial \tau_j} - B^N_{\tau_j}.$$

Using the definitions of the B_{τ_j} above, it is straightforward to write $\frac{\partial B_{\tau_m}}{\partial \tau_j}$ as a function of the elasticities of the variables K_E, K and N with respect to taxes. For instance:

$$\begin{aligned}
(1-\tau_K)\frac{\partial B_{\tau_N}}{\partial \tau_K} &= wN\overline{e}^N_{\tau_K} \\
(1-\tau_K)\frac{\partial B_{\tau_K}}{\partial \tau_K} &= -\pi_F N\overline{e}^N_{\tau_K} + (r_E - r_F)K_E\overline{e}^{K_E}_{\tau_K} + (r_F-\delta)(K\overline{e}^K_{\tau_K}) \\
(1-\tau_K)\frac{\partial B_{\tau_W}}{\partial \tau_K} &= K\overline{e}^K_{\tau_K} \\
(1-\tau_K)\frac{\partial B_{\tau_C}}{\partial \tau_K} &= (1-\tau_K)\frac{\partial}{\partial \tau_K}(Y-\delta K-\overline{G}) = (1-\tau_K)\frac{\partial B_{\tau_K}}{\partial \tau_K} + (1-\tau_K)\frac{\partial B_{\tau_N}}{\partial \tau_K}.
\end{aligned}$$

The expressions, for each j, for $\frac{\partial B_{\tau_j}}{\partial \tau_W}$ are identical to these, except that the elasticities on the right hand side are all replaced by elasticities with respect to τ_W.

In the next section, we show that the optimal tax rates can be written as functions of these derivatives of the B_{τ_j} terms. Therefore, the results above allow the derivatives of the B_{τ_j} to be calculated, given the elasticities of K_E, K and N with respect to taxes, which can be used to calculate the optimal tax rates.

4.2 Characterizing Optimal Taxes

We next characterize the optimal taxes chosen by the planner. We focus on an optimal steady state tax policy. In particular, we assume that the planner chooses steady state tax rates $\{\tau^*_K, \tau^*_W, \tau^*_N\}$, and an aggregate steady state $\mathscr{S}$ of the economy consistent with these tax rates, in order to maximize our welfare measure – the present discounted utility of a newborn agent in the steady state.

Since we have already established the marginal effects of tax changes on welfare, it is relatively straightforward to derive the optimal tax rates. The first order conditions for the optimal choice of τ_K and τ_W is simply that $\frac{\partial \Delta^N}{\partial \tau_j} = 0$ for these three tax rates. The optimal choice of τ_N can then be inferred from government budget balance. Specifically, the first order condition for each τ_j is

$$0 = (1 + A_{\tau_j}\tau_j)(1-\tau_N)wN\frac{\partial \Delta^N}{\partial \tau_j} = (1 + A_j\tau_j)(B_{\tau_j} - B^N_{\tau_j}) + \sum_{m\in\{K;W;C;N\}} \tau_m B_{\tau_m}\bar{e}^{B_{\tau_m}}_{\tau_j},$$

where $\bar{e}^{B_{\tau_m}}_{\tau_j}$ denotes the general equilibrium elasticity of B_{τ_m} with respect to τ_j, and $A_{\tau_K} = -1$ and $A_{\tau_W} = 0$.

Using the government budget constraint we can eliminate τ_N from the first order condition for τ_j and obtain:

$$0 = (1 + A_j\tau_j)(B_{\tau_j} - B^N_{\tau_j}) + \overline{G}\bar{e}^{B_{\tau_N}}_{\tau_j} + \sum_{m\in\{K;W;C\}} \tau_m B_{\tau_m}(\bar{e}^{B_{\tau_m}}_{\tau_j} - \bar{e}^{B_{\tau_N}}_{\tau_j}).$$

Since $B_{\tau_N} = wN$ it follows that $\bar{e}^{B_{\tau_N}}_{\tau_j} \equiv \bar{e}^N_{\tau_j}$, since the elasticities are calculated holding w constant.

The first order conditions can be compactly written in matrix form as follows:

$$\mathbf{0} = (B - B^N)\mathbf{1} + A(B - B^N)\mathcal{T} + \overline{G}\bar{\mathbf{e}}^N + \mathcal{E}B\mathcal{T} + B_{\tau_C}\tau_C\bar{\mathbf{e}}^{\mathbf{B}_{\tau_\mathbf{C}}} - \bar{\mathbf{e}}^N\mathbf{1}^T B\mathcal{T},$$

where $\mathbf{1}$ denotes the column vector $(1,1)^T$, $\bar{\mathbf{e}}^N$ denotes the column vector $(\bar{e}^N_{\tau_K}, \bar{e}^N_{\tau_W})^T$, $\mathcal{T}$ denotes the column vector $(\tau_K, \tau_W)^T$, $\bar{\mathbf{e}}^{\mathbf{B}_{\tau_\mathbf{C}}}$ denotes the column vector $(\bar{e}^{B_{\tau_C}}_{\tau_K}, \bar{e}^{B_{\tau_C}}_{\tau_W})^T$ and A, B, B^N and $\mathcal{E}$ are defined as follows:

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} B_{\tau_K} & 0 \\ 0 & B_{\tau_W} \end{pmatrix},$$

$$B^N = \begin{pmatrix} B^N_{\tau_K} & 0 \\ 0 & B^N_{\tau_W} \end{pmatrix}, \qquad \mathcal{E} = \begin{pmatrix} \bar{e}^{B_{\tau_K}}_{\tau_K} & \bar{e}^{B_{\tau_W}}_{\tau_K} \\ \bar{e}^{B_{\tau_K}}_{\tau_W} & \bar{e}^{B_{\tau_W}}_{\tau_W} \end{pmatrix}.$$

Rearranging the first order condition, we therefore obtain the solution for the optimal tax vector:

$$\mathcal{T} = -\left(A\left(B - B^N\right) + \mathcal{E}B - \bar{\mathbf{e}}^N\mathbf{1}^T B\right)^{-1}\left(\left(B - B^N\right)\mathbf{1} + \overline{G}\bar{\mathbf{e}}^N + B_{\tau_C}\tau_C\bar{\mathbf{e}}^{\mathbf{B}_{\tau_\mathbf{C}}}\right). \tag{28}$$

Thus, optimal taxes are entirely a function of three components – the size of the tax base of each tax B, the elasticity of each tax base with respect to tax rates ($\mathcal{E}$ and $\bar{\mathbf{e}}^N$), and the degree to which each tax falls upon workers, $\mathbf{B}^N$. The latter is relevant because the utilitarian planner has a motive to redistribute to workers if they have lower average consumption than entrepreneurs in the steady state. Since the elasticities $\mathcal{E}$ shows up in the inverse term in equation (28), higher values of the elasticities will tend to make the right hand side of the equation smaller, and drive optimal taxes on capital and wealth towards zero. This is consistent with standard optimal tax principles, that it is optimal to set lower a lower tax rate on a particular activity if the associated tax base is more elastic with respect to the tax rate.

The derivation of the optimal tax formula here did not rely on most of the specific assumptions of the model. The optimal tax formula above would be identical if different functional forms were chosen for utility, or the distribution of entrepreneurial ability or entrepreneurial shocks, or if the endogenous financial friction that we model were replaced by an exogenous constraint on entrepreneur's ability to issue equity and debt. However, an advantage of the specific model assumptions is that they allow the elasticities in equation (28) to be computed relatively easily, using the comparative static results in the previous section. In the next section, we use these elasticities to calculate optimal taxes for specific values of the parameters.

4.3 A Numerical Calibration

To interpret the magnitudes of these optimal taxes in practice, we undertake a numerical calibration. To that end, we construct a benchmark economy that has the same primitives as the economy outlined in Section 2 and in which we set taxes at their current levels in the United States. We calibrate the benchmark economy at annual frequency and summarize parameter values in Table 1.

Demographics We set the mortality rate γ to 1%, corresponding to a life expectancy of 100 years. We choose the discount rate ρ so that in the steady state the average return to capital, weighted by capital shares and net of depreciation, is 4% (McGrattan and Prescott, 2001).

Technology We set the depreciation rate δ to 7%, approximately the average depreciation rate in the US fixed asset tables, and the autocorrelation of the productivity shock $1-\lambda_\theta$ to 0.885, as in Cooper and Haltiwanger (2006). Following Panousi (2012), we calibrate φ to 0.15, to match the variance of entrepreneurial capital. We choose the lowest possible realization of the idiosyncratic productivity shock, $\underline{\epsilon}$, to match a debt-to-asset ratio for entrepreneurs of 0.35 (Mehrotra and Crouzet, 2017, Boar and Midrigan, 2019). Lastly, we assume that the final output production technology is Cobb-Douglas $Y = Y_E^{\alpha_E} Y_F^{\alpha_F} N^{1-\alpha_N}$ and calibrate α_E, α_F, α_N and π, the probability of a high type (i.e. $\theta = 1$), to match a labor share of 2/3, the share of households who are entrepreneurs, the return to risky capital and the risk-free interest rate. We target an average return to capital in risky projects gross of depreciation of 15% so that the return net of depreciation is 8%, approximately the annual rate of return to equities in the US over the twentieth century (Mehra and Prescott, 2003). We calibrate to a risk-free return gross of depreciation of 8%, consistent with a risk-free interest rate of 1% which is close to the average return of relatively riskless securities in the US over the twentieth century (Mehra and Prescott, 2003). Finally, we target a share of entrepreneurs of 11.7%, as reported by Boar and Midrigan (2020) using data from the Survey of Consumer Finances.

Financial friction Since the entrepreneur's optimal contract is equivalent to an equity and debt contract, we calibrate the parameter ϕ, indicating the severity of the agency friction, to 0.76 so as to match the equity share of business owners in the US data. We use the Survey of Consumer Finances (National) Survey of Small Business Finances to document that entrepreneurs own, on average, 84% of their firm's equity.[14]

Tax system We set the consumption tax τ_C to 11%, following Altig et al. (2001) and Cagetti and Nardi (2009). We set τ_E to 20%, in line with the US corporate tax rate for small businesses reported in the OECD Tax Database, and τ_W to zero, in line with the current practice in the US.[15] We choose $\bar{G}$, so that the share of government spending is 20% of GDP, the historical average in the US over the past four decades.

Optimal Taxes. To quantify the various mechanisms, we consider separately how optimal taxes depend on partial and general equilibrium effects in this calibrated model. The results are shown in Table 2 below. The fourth row of the table shows the correct optimal taxes in the calibrated model. The first three rows show the optimal taxes that arise when various factors are held constant. We discuss each row in turn.

The first row calculates optimal taxes when the fraction of workers N is held fixed at the initial steady state value and general equilibrium effects are ignored. These optimal taxes are inferred by calculating the terms in the optimal tax equation (28) above in the initial steady state, but using the partial equilibrium elasticities $e_{\tau_j}^{B_{\tau_i}}$ instead of the general equilibrium elasticities $\overline{e}_{\tau_j}^{B_{\tau_i}}$ and setting $\overline{e}_{\tau_j}^{N} = 0$. We see that, in this case, optimal taxes on capital income are slightly negative and optimal wealth taxes essentially zero. The reason for this is that the tax base for both capital income and wealth taxes is highly elastic with respect to tax rates, since, as discussed in the previous section, the capital stock is highly elastic with respect to tax changes. As a consequence of this high elasticity, tax revenue is maximized at tax rates close to zero (since raising the tax seriously reduces the size of the tax base). The optimal

[14] See Appendix C for a detailed discussion of our treatment of the data.

[15] The calibrated value of τ_K is an average of the tax rate between 2000 and 2016 and includes both federal and state taxes.

Table 1: Parameter Values

Parameter	Value used	Target moment
γ	0.010	Lifespan 100 Years
ρ	0.009	Average net return to capital 4%
δ	0.070	Depreciation
λ_θ	0.115	Profitability autocor. (Cooper and Haltiwanger, 2006)
φ	0.150	Small Bus. Risk (Panousi, 2012)
$\underline{\epsilon}$	0.350	Debt-to-asset ratio (Boar and Midrigan, 2019)
α_E	0.193	Labor share 2/3
α_F	0.137	Risk-free rate
α_N	0.593	Fraction of entrepreneurs (Boar and Midrigan, 2020)
π	0.015	Return to Equity
τ_C	0.110	Consumption tax rate (Cagetti and Nardi, 2009)
τ_K	0.200	Corp. tax rate small businesses (OECD Tax Database)
τ_W	0	Current US level
$\bar{G}$	0.200	Govt. spending/GDP
ϕ	0.757	Small Bus. Owner Equity Share (SSBF)

capital income tax is actually slightly negative, since negative capital income taxes increase K_E. That is, negative capital income taxes not only increase aggregate capital accumulation, but also increase the fraction of capital in the risky technology, by raising the relative return to the risky technology and redistributing resources to high ability entrepreneurs, who earn a higher rate of return to capital and therefore disproportionally benefit from lower capital income taxes. Through these channels, negative capital income taxes increase steady state output and therefore the revenue earned by the government from labor income taxes, making them desirable.

The second row of Table 2 shows optimal taxes allowing for general equilibrium effects on prices, but holding fixed N and all elasticities. These optimal taxes are calculated as in the first row of the table, except that general equilibrium elasticities are used, but with $\overline{e}^N_{\tau_j}$ set to zero.[16] Here, we see very different optimal taxes from the partial equilibrium case. As explained above, the only effect of a general equilibrium price changes that is relevant for optimal taxes is the effect of changes in $\mathcal{R}$. Higher rates of capital taxation are much more desirable in the general equilibrium case because they increase $\mathcal{R}$. To understand this mechanism, recall that $\mathcal{R}$ reflects the degree to which entrepreneurs' profits represent returns

[16] As such, all other general equilibrium elasticities are also slightly different from the full general equilibrium model, because we ignore all changes in K, K_E and other aggregate variables that are due to changes on N, since N is held constant.

on investing in the risky technology, rather than economic rent from selling final goods. As discussed above, $\mathcal{R}$ is decreasing in K_E (near the steady state, where $\mathcal{R} = 0$), because higher K_E reduces the pre-tax rate of return on the risky technology. Consequently, a higher value of capital income taxes reduces K_E (by reducing the post-tax return to the risky technology) but this in turn raises $\mathcal{R}$. The rise in $\mathcal{R}$ redistributes profits to those who invest in the risky technology, and so increases the return to that technology, which is $\tilde{r}_X(1 + \mathcal{R})$. The rise in the return to the risky technology substantially mitigates the fall in K_E due to rising capital income taxes. Consequently, allowing for these general equilibrium effects, the elasticity of capital income to the tax rate is substantially lower than it was in partial equilibrium, and so the optimal capital income tax is higher. The high value of optimal capital income taxes encourages a slightly negative tax on wealth, since this encourages capital accumulation and raises the revenue earned by the capital income taxes.

The third row of Table 2 repeats the same analysis, but fully allows for the variation of N in response to changes in tax rates. As such, optimal taxes are calculated using the optimal tax equation (28) above and using the correct general equilibrium elasticities calculated in the initial steady state. Allowing for variation in N reduces the optimal tax rate on capital income significantly, because high capital income taxes reduce entry into entrepreneurship, which reduces output and therefore the government's tax revenue.

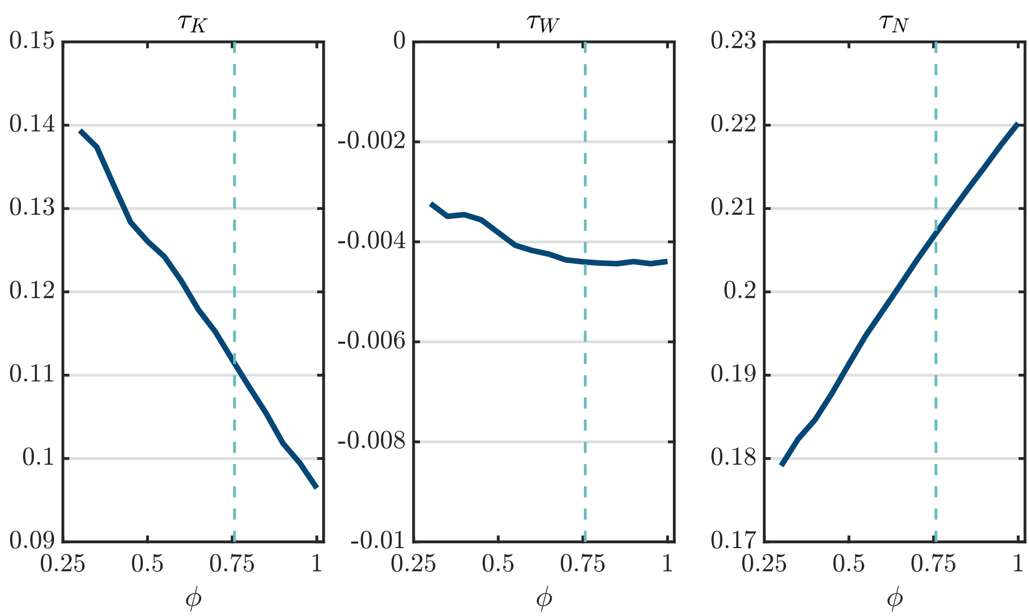
The optimal taxes on the third row are not quite numerically correct, because changing tax rates in turn affects the elasticities in the optimal tax equation (28). Therefore, to correctly calculate optimal tax rates, it is necessary to repeatedly resolve the equation in an iterative process – the equation is first used to calculate conjectured optimal tax rates, then the elasticities are recalculated at the new tax rates, and the equation is solved again yielding a new set of tax rates and so on. Comparing row 4 and row 3 of the table reveals that this iterative process produces optimal taxes that are very close to those obtained by just applying equation (28) once. This reveals that a very good approximation to the optimal taxes is obtained, even if we only know the elasticities in the initial steady state. An advantage of this is that elasticities in the initial steady state are more easily estimated empirically.

Table 2: Optimal Taxes in Partial and General Equilibrium

	τ_K^*	τ_W^*	τ_N^*
Partial Equilibrium, Fixed N	−0.9%	0%	23.4%
General Equilibrium, Fixed N	57.1%	−2%	14.1%
General Equilibrium	9.6%	−0.4%	21.8%
General Equilibrium, Iterative	11.1%	−0.4%	20.7%

The Importance of Financial Frictions. We next analyze the implications of financial frictions for optimal taxes. To that end, Figure 1 shows how optimal taxes vary with the severity of financial frictions, governed by the parameter ϕ. To facilitate the interpretation of the units on the horizontal axis, we note that varying ϕ from 0.3 (mild financial friction) to 1 (severe financial friction) implies that the share of equity that entrepreneurs own in their business increases from 33% to 110%. The dashed vertical line marks the value of ϕ in our calibration. At the calibrated ϕ the optimal tax rates on capital income, wealth and labor income are 11.14%, -0.44% and 20.7%, respectively. That is, it is optimal to tax capital income and subsidize wealth accumulation. Three features are immediately evident from the figure. First, that optimal wealth taxes are always close to zero for all values of ϕ we consider. Second, that optimal capital income taxes are positive for all values of ϕ, and third, optimal capital income taxes are decreasing in the value of ϕ. We discuss each of these three in turn.

Figure 1: Optimal Taxes and Financial Frictions



First, the optimal wealth tax is close to zero, regardless of the value of ϕ. This arises from our optimal taxation formula because the elasticity of all aggregate variables with respect to the wealth tax is very large, as was also discussed above in Section 3.4.2. As a consequence of this, the elasticities of all tax bases with respect to the wealth tax rate are large, leading to an optimal wealth tax close to zero. The optimal wealth tax is generally found to be slightly negative, because a negative wealth tax increases capital accumulation, which raises the government's revenue from the capital income tax.

Second, the optimal capital income tax rate is positive for most values of ϕ. As discussed on page 41, this is because positive capital income taxes raise $\mathcal{R}$, which redistributes resources to the most efficient entrepreneurs and raises the return to capital. As such the increase in $\mathcal{R}$ tends to increase K and K_E, which therefore mitigates the negative effects of capital income taxes on capital accumulation and on the allocation of capital.

Finally, it is evident from Figure 1 that the optimal capital income tax is decreasing in ϕ. This is because, when ϕ is lower, financial frictions are tighter and so less capital is allocated to the risky technology in the initial steady state. Therefore, with lower ϕ, the return on capital in the risky technology is relatively higher. This means that a cut in capital income taxes will increase the size of the tax bases by relatively more, encouraging the planner to set a lower level of capital income tax to increase tax revenue.

Sensitivity analysis. Lastly, we explore how our results depend on values of other relevant parameters. First, as is well known in the public finance literature, optimal taxes depend on the underlying distribution of ability. In our case, however, variation in the dynamic process governing entrepreneurial ability appears to matter little to results. To illustrate, we consider two alternative calibrations of the autocorrelation of entrepreneurial ability, which is governed by the parameter λ_θ. In the first calibration we reduce λ_θ from 0.115 to 0.05, corresponding to a more persistent process for entrepreneurial ability. In the second one, we increase this parameter to 0.2, indicative of a more persistent process for entrepreneurial ability. In the first case, we find that optimal taxes on capital income, wealth and labor income are equal to 13.9%, -0.41% and 18.7%, respectively. In the second case, these are equal to 9.05%, -0.39% and 22.5%, respectively.

On the other hand, we find that results are more sensitive to the parameters governing the returns to scale – optimal taxes on capital income decrease as parameter values move towards constant returns to scale. Again, we consider two alternative calibrations of the model. In the first calibration, we set $\alpha_N = 0.5$, implying that an elasticity of returns to scale of 0.83, which is lower than the benchmark calibration's 0.92 (where 1 denotes constant returns). We find that, with $\alpha_N = 0.5$, optimal taxes on capital income, wealth and labor income are equal to 21.27%, -0.41% and 27%, respectively. In the second calibration, we set $\alpha_N = 0.65$, implying an elasticity of returns to scale of 0.99 – almost constant returns. In this case, we find that the optimal tax on capital income, wealth and labor income are -4.14%, -0.21% and 14.86%. Thus, moving towards constant returns strongly reduces optimal taxes on capital income, and slightly reduces optimal taxes on labor income. This is because, as discussed above, a major force influencing optimal taxes is the effect of capital income taxes on $\mathcal{R}$, which governs how far entrepreneurial profits are due to sales of the final good, and how far they are due to the returns on risky capital. When parameters move towards constant returns, entrepreneurs' profits from selling the final good fall to zero, regardless of the level of taxation, and so the mechanism through changes in $\mathcal{R}$ disappears. Consequently, optimal taxes on capital income become similar to those in the partial equilibrium case in Table 2, because the general equilibrium effect through $\mathcal{R}$ disappears. As explained above, in the partial equilibrium (or close to constant returns) case optimal taxes on capital income are slightly negative, because this increases capital accumulation and improves the efficiency of the allocation of capital, which raises output and therefore government tax revenue.

5 Conclusion

We examine optimal linear taxation in a setting with endogenous entry and financial frictions. Financial frictions imply that the distribution of wealth across entrepreneurs with different productivity levels affects how efficiently capital is allocated in the economy – a force missing from models without financial frictions. The planner chooses taxes on capital income, wealth and labor income to maximize the steady state welfare of a newborn agent. In the model, newborn agents decide whether to become workers or entrepreneurs. Workers supply labor inelastically, while entrepreneurs operate a production technology that uses capital and are subject to a financial constraint that stems from informational frictions. As in the data, entrepreneurs are relatively richer, leading to a redistributive motive for capital income and wealth taxation.

Our model is analytically tractable and we characterize optimal steady state taxes as closed-form functions of the size of tax bases and the elasticity of tax bases with respect to taxes, in the tradition of the "sufficient statistics" approach to optimal taxation. Evaluated at empirically plausible parameters, we that it is optimal to tax wealth at a rate close to zero and to tax capital income at a positive rate, albeit lower than the labor income tax rate. This result reflects a large elasticity of the wealth tax base with respect to the wealth tax and a somewhat smaller elasticity of the capital income tax base to the capital income tax. The latter is smaller because capital income taxes lead to price changes in general equilibrium that raise the return to risky investment and redistribute resources from low ability to high ability entrepreneurs, which partially counteracts the otherwise negative effects of capital income taxes on capital accumulation.

Our optimal tax formula is general and does not depend on many of the details of the model, such as functional form assumptions or the specific details of the financial friction. However, to maintain analytical tractability and for simplicity of exposition, we restrict our analysis to linear taxes and abstract from features of potential relevance such as worker heterogeneity, elastic labor supply, bequests or aging. Relaxing these assumptions is left for future work and while these considerations will doubtlessly affect the value of optimal taxes, the channels we emphasize in this paper remain operative.

References

Abo-Zaid, Salem. 2014. "Optimal capital-income taxation in a model with credit frictions." *The B.E. Journal of Macroeconomics* 14(1):147 – 172.

Altig, David, Alan J. Auerbach, Laurence J. Kotlikoff, Kent A. Smetters and Jan Walliser. 2001. "Simulating Fundamental Tax Reform in the United States." *The American Economic Review* 91(3):574–595.

Basseto, Marco and Wei Cui. 2020. A Ramsey Theory of Financial Distortions. Working paper.

Benhabib, Jess and Bálint Szőke. 2019. Optimal Positive Capital Taxes at Interior Steady States. Working Paper 25895 National Bureau of Economic Research.

Bernanke, Ben S, Mark Gertler and Simon Gilchrist. 1999. "The financial accelerator in a quantitative business cycle framework." *Handbook of macroeconomics* 1:1341–1393.

Biljanovska, Nina. 2019. "Optimal Policy In Collateral Constrained Economies." *Macroeconomic Dynamics* 23(2):798–836.

Biljanovska, Nina and Alexandros P. Vardoulakis. 2019. "Capital taxation with heterogeneous discounting and collateralized borrowing." *Journal of Macroeconomics* 60:97 – 109.

Boar, Corina and Virgiliu Midrigan. 2019. Markups and Inequality. Working Paper 25952 National Bureau of Economic Research.

Boar, Corina and Virgiliu Midrigan. 2020. Efficient Redistribution. Working Paper 27622 National Bureau of Economic Research.

Cagetti, Marco and Mariacristina De Nardi. 2009. "Estate Taxation, Entrepreneurship, and Wealth." *The American Economic Review* 99(1):85–111.

Cagetti, Marco and Mariacristinasc De Nardi. 2009. "Estate Taxation, Entrepreneurship, and Wealth." *American Economic Review* 99(1):85–111.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica* 54(3):607–622.

Chari, V. V. and Patrick J. Kehoe. 1999. Optimal Fiscal and Monetary Policy. Working Paper 6891 National Bureau of Economic Research.

Chen, Yunmin, YiLi Chen and C.C. Yang. 2019. Implementing the Modified Golden Rule? Optimal Ramsey Taxation with Incomplete Markets Revisited. Working paper.

Chetty, Raj. 2008. Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. Working Paper 14399 National Bureau of Economic Research.

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *The Review of Economic Studies* 73(3):611–633.

Cullen, Julie Berry and Roger H Gordon. 2007. "Taxes and entrepreneurial risk-taking: Theory and evidence for the US." *Journal of Public Economics* 91(7-8):1479–1505.

Devereux, Michael. 2009. Taxing Risky Investment. Technical report Oxford University Centre for Business Taxation.

Evans, David. 2015. Optimal Taxation with Persistent Idiosyncratic Investment Risk. Working paper.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo and Daphne Chen. 2019. Use It or Lose It: Efficiency Gains from Wealth Taxation. Working paper.

Hayek, Friedrich August. 1960. *The Constitution of Liberty.(Second Impression.)*. University of Chicago Press.

Itskhoki, Oleg and Benjamin Moll. 2019. "Optimal Development Policies With Financial Frictions." *Econometrica* 87(1):139–173.

Judd, Kenneth L. 1985. "Redistributive taxation in a simple perfect foresight model." *Journal of Public Economics* 28(1):59 – 83.

Kitao, Sagiri. 2008. "Entrepreneurship, taxation and capital investment." *Review of Economic Dynamics* 11(1):44–69.

McGrattan, Ellen R and Edward C Prescott. 2001. Is the Stock Market Overvalued? Working Paper 8077 National Bureau of Economic Research.

Mehra, Rajnish and Edward C Prescott. 2003. "The equity premium in retrospect." *Handbook of the Economics of Finance* 1:889–938.

Mehrotra, Neil and Nicolas Crouzet. 2017. Small and Large Firms over the Business Cycle. Working paper.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and misallocation: Evidence from plant-level data." *American economic review* 104(2):422–58.

Panousi, Vasia. 2012. "Capital taxation with entrepreneurial risk." *forthcoming Journal of Monetary Economics* .

Panousi, Vasia and Catarina Reis. 2014. "Optimal capital taxation with idiosyncratic investment risk." *Working paper* .

Panousi, Vasia and Catarina Reis. 2019. A Unified Framework for Optimal Taxation with Undiversifiable Risk. Working paper.

Phelan, Tom. 2019. On the Optimality of Differential Asset Taxation. Working papers Federal Reserve Bank of Cleveland.

Piketty, Thomas and Emmanuel Saez. 2013. "A theory of optimal inheritance taxation." *Econometrica* 81(5):1851–1886.

Piketty, Thomas, Emmanuel Saez and Stefanie Stantcheva. 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy* 6(1):230–71.

Rotberg, Shahar and Joseph B. Steinberg. 2019. Tax the Wealthy? A Dynamic General Equilibrium Analysis of Progressive Wealth Taxation. Working paper.

Saez, Emmanuel and Gabriel Zucman. 2019. "Progressive wealth taxation." *Brookings Papers on Economic Activity* .

Saez, Emmanuel and Stefanie Stantcheva. 2018. "A simpler theory of optimal capital taxation." *Journal of Public Economics* 162:120 – 142.

Shourideh, Ali. 2014. Optimal Taxation of Wealthy Individuals. Working paper Wharton School.

Smith, Matt, Owen Zidar and Eric Zwick. 2020. Top Wealth in America: New Estimates and Implications for Taxing the Rich. Working paper.

Straub, Ludwig and Ivan Werning. 2014. Positive Long Run Capital Taxation: Chamley-Judd Revisited. Working Paper 20441 National Bureau of Economic Research.

Varian, Hal R. 2014. *Intermediate microeconomics with calculus: a modern approach*. WW Norton & Company.

Appendices

A Discrete Time Model

A.1 Worker's Optimization Problem

We derive the solution to the worker's problem. As a first step, we rewrite the worker's problem, replacing the state variable a^N by P^N.

Observe that the worker's budget constraint may be rewritten as:

$$c_t^N + \frac{R_{F,t}}{1+\tau_{C,t}} \frac{(1-\gamma)\, P_{t+1}^N}{R_{F,t}} = \frac{R_{F,t}}{1+\tau_{C,t}} P_t^N.$$

Hence, abusing notation, the worker problem may be written as:

$$V^N(P^N, X) = \max_{c\geq 0, P'\geq 0} \left(\log\left(c^N\right) + (1-\rho)\,(1-\gamma)\, V^N\left(P^{N\prime}, X'\right)\right)$$

$$\text{s.t.} \quad c^N + \frac{1-\gamma}{1+\tau_C} P^{N\prime} = \frac{R_F}{1+\tau_C} P^N$$

We shall solve the problem by guess and verify. We guess:

$$V^N(P^N, X) = Q_t + \frac{1}{1-(1-\rho)(1-\gamma)} \log\left(P_t^N\right)$$

Taking first order conditions, we get:

$$\left(\frac{R_F}{1+\tau_C} P^N - \frac{1-\gamma}{1+\tau_C} P^{N\prime} \right)^{-1} \frac{1-\gamma}{1+\tau_C} = (1-\rho)(1-\gamma) \frac{\partial V^N(P^{N\prime}, X')}{\partial P^{N\prime}}$$

Using the guess and simplifying, we obtain the worker's policy functions:

$$P^{N\prime} = (1-\rho)\,(1-\gamma)\, \frac{R_F}{1-\gamma} P^N$$

$$c^N = \left[1-(1-\rho)\,(1-\gamma)\right] \frac{R_F}{1+\tau_C} P^N$$

To verify the guess, we plug the policy function conjectures back into the conjectured value function:

$$\begin{aligned}
V^N(P^N,X) =& \max\left(\log\left([1-(1-\rho)(1-\gamma)]\frac{R_F}{1+\tau_C}P^N\right)\right.\\
&\left.+(1-\rho)(1-\gamma)\left(Q'+\frac{1}{1-(1-\rho)(1-\gamma)}\log\left((1-\rho)(1-\gamma)\frac{R_F}{1-\gamma}P^N\right)\right)\right)\\
V^N(P^N,X) =& \log\left((1-(1-\rho)(1-\gamma))\frac{R_F}{1+\tau_c}\right)\\
&+(1-\rho)(1-\gamma)\left(Q'+\frac{1}{1-(1-\rho)(1-\gamma)}\log\left((1-\rho)(1-\gamma)\frac{R_F}{1-\gamma}\right)\right)\\
&+\frac{1}{1-(1-\rho)(1-\gamma)}\log P^N\\
\implies \quad V^N(P^N,X) =& \underbrace{V^N(1,X)}_{=Q}+\frac{1}{1-(1-\rho)(1-\gamma)}\log P^N
\end{aligned}$$

□

A.2 Proof of Lemma 1

Note that the technology associated with an entrepreneur's risky and risk-free projects displays constant returns to scale. Consider an entrepreneur Alice, who at some period t has $q > 0$ times as much lifetime resources P as an entrepreneur Bob. The constant returns to scale properties imply that Alice can produce q times as much of each intermediate good as Bob, for each ϵ and θ and consume q times as much each period. Since, for any c, $\log(qc) \equiv \log(q) + \log(c)$, Alice's present discounted utility from these choices would then be the same as Bob's plus an additional $\sum_{j=0}^{\infty}(1-\rho)^j(1-\gamma)^j \log(q) = \frac{\log(q)}{1-(1-\rho)(1-\gamma)}$. Since these choices are possible for Alice, it must be that $V(qP,\theta,X) \geq \frac{\log(q)}{1-(1-\rho)(1-\gamma)} + V(P,\theta,X)$. However, on the other hand, Bob can choose to do everything that Alice does only $\frac{1}{q}$ times as much. By the same logic as before, doing so would yield Bob a present discounted utility equal to Alice's minus $\frac{\log(q)}{1-(1-\rho)(1-\gamma)}$. Therefore, it must be the case that $V(P,\theta,X) \geq V(qP,\theta,X) - \frac{\log(q)}{1-(1-\rho)(1-\gamma)}$. Comparing these two inequalities that V must fulfil, it is immediate that it cannot satisfy both unless $V(qP,\theta,X) = \frac{\log(q)}{1-(1-\rho)(1-\gamma)} + V(P,\theta,X)$. In that case, it must be that $V(P,\theta,X) \equiv \frac{\log(P)}{1-(1-\rho)(1-\gamma)} + V(1,\theta,X)$. Let $\overline{V}(\theta,X)$ denote $V(1,\theta,X)$. Then it follows that $V(P,\theta,X) = \overline{V}(\theta,X) + \frac{\log(P)}{1-(1-\rho)(1-\gamma)}$. □

A.3 Proof of Lemma 2

Suppose that the entrepreneur is sells fraction $1-\frac{\phi(1+\tau_C)}{(1-\tau_K)(1-\delta)}$ equity in her risky project, at a price of 1 per unit of k_E invested in the project, retaining the rest herself. Suppose further, that she issues b so that $R_F b$ is equal to the total end of period value of her risky project

$\epsilon = \underline{\epsilon}$, including the profits earned from this project. Then,

$$R_F b = (((\underline{\epsilon} - 1)(1-\delta) + \theta r_E)(1-\tau_K) - \tau_W)k_E$$

and

$$\hat{b} = R_F b + \left(1 - \frac{\phi(1+\tau_C)}{(1-\tau_K)(1-\delta)}\right)(\epsilon - \underline{\epsilon})(1-\tau_K)k_E$$

Substituting these into the entrepreneur's budget constraints, and using the definitions of ω, and $\underline{\omega}$, reveals that, in every state of the world, this contract would give the entrepreneur the same end of period resources ω as the optimal contract.

□

A.4 Proof of Lemma 3

In an equilibrium of the economy, it must be the case every period that:

$$-\phi + (r_E - r_F)\frac{1-\tau_K}{1+\tau_C} \quad \leq \quad 0 \tag{A.1}$$

If this were not the case, then equations (9) and the expression for $\underline{\omega}$

$$\underline{\omega} = \left(-\phi + (r_E - r_F)\frac{1-\tau_K}{1+\tau_C}\right)k_E + \left(\frac{1}{1+\tau_C}R_F\right)P$$

imply that ω is strictly increasing in k_E for all $\epsilon > 0$ for entrepreneurs with $\theta = 1$. Then, by choosing an arbitrarily high k_E, such entrepreneurs will be able to achieve arbitrarily high consumption and therefore utility. Therefore, if this condition held, some entrepreneurs would desire to allocate an infinite amount of capital to their risky projects, which cannot be an equilibrium, since the capital stock is finite each period. □

B Continuous Time Model

B.1 Environment and Equilibrium with Period Length Δ

Let $\Delta \in (0,1]$ denote the length of a period. Over a period of length Δ, agents value future consumption at rate $(1-\rho\Delta)$, die with probability $\gamma\Delta$, capital depreciates at rate $\delta\Delta$, entrepreneurs draw a new productivity θ with probability $\lambda_\theta\Delta$ and shocks ϵ satisfy

$$\epsilon = \underline{\epsilon} + (1-\underline{\epsilon})\exp\left(\varphi\sqrt{\Delta}\cdot\mathcal{R} - \frac{\varphi^2\Delta}{2}\right), \tag{B.1}$$

where $\mathcal{R} \sim N(0,1)$ and φ is a parameter determining the variance of ϵ. This assumption implies that $E[\epsilon] = 1$ and $\text{Var}(\log(\epsilon - \underline{\epsilon})) = \varphi^2 \Delta$. Risky projects produce $\theta \epsilon k_E \Delta$ and risk-free projects produce $k_F \Delta$.

An entrepreneur values her expected future consumption stream according to $\mathbb{E}_t \left[\sum_{j=0}^{\infty} (1-\rho\Delta)^j (1-\gamma\Delta)^j \log(c_{i,t+j\Delta}) \Delta \right]$. We define the lifetime income of an entrepreneur who does not produce any intermediate goods and lends all of his endowment to the bank at riskfree rate $1 + (R_F - 1)\Delta = 1 + \tilde{R}_F \Delta$ as,

$$P_{i,t} := a_{i,t} + \underbrace{\sum_{j=0}^{\infty} \left[\frac{\pi^{\star}_{t+j} \Delta (1 - \tau_{K,t+j})(1-\gamma\Delta)^j}{\Pi_{k=0}^{j} 1 + \tilde{R}_{F,t+k} \Delta} \right]}_{=:F_t}. \tag{B.2}$$

A worker's preferences are described by the lifetime utility function $\sum_{j=0}^{\infty} (1-\rho\Delta)^j (1-\gamma\Delta)^j \log\left(c^N_{t+j\Delta}\right) \Delta$. The government sets taxes τ_C, τ_N, τ_K and $\tau_W \Delta$ per period, and has to finance exogenous expenditure $\overline{G}\Delta$.

First, consider the worker's problem for a given period length Δ:

$$V^N(a^N, X) = \max\left(\log\left(c^N\right)\Delta + (1-\rho\Delta)(1-\gamma\Delta) V^N\left(a^{N\prime}, X'\right)\right) \tag{B.3}$$

subject to:

$$c^N \Delta (1 + \tau_C) + (1-\gamma\Delta) a^{N\prime} = w\Delta(1 - \tau_N) + (1 + (R_F - 1)\Delta) a^N,$$

and non-negativity constraints on c^N, $a^{N\prime}$.

Second, the entrepreneur's problem is to solve:

$$\begin{aligned} V(P,\theta,X) \;&=\; \sup \int_{\epsilon>0} \Bigg(\log(c)\Delta + (1-\rho\Delta)(1-\gamma\Delta) \\ &\quad \times\; E\left[V(P',\theta',X') \middle| \theta \right] \Bigg) \mathrm{d}H(\epsilon), \end{aligned} \tag{B.4}$$

subject to:

$$\begin{aligned} & c\Delta + \frac{1-\gamma\Delta}{1+\tau_C} P' = \omega, \\ & \quad (1+\tau_C)\,\omega = (1+\tau_C)\, c_H \Delta - \hat{b} + (1-\tau_K)\,\pi\Delta + (\tau_K \delta\Delta - \tau_W \Delta)\, k + (1-\delta\Delta)k + (1-\gamma\Delta)F', \\ & \qquad \pi\Delta = (r_E \Delta y_E + r_F \Delta y_F) + (y\Delta - w\Delta n - r_E \Delta y_E^d - r_F \Delta y_F^d) + (1-\delta\Delta)\left((\varepsilon - 1)k_E - k_H\right), \end{aligned}$$

the production functions $c_H = \phi k_H$, $y_E = \theta\epsilon k_E$, $y_F = k_F$, $y = f(y_E^d, y_F^d, n)$, the incentive compatibility constraint:

$$\frac{(1-\tau_K)\left(r_E\Delta+(1-\delta\Delta)\right)k_E}{1+\tau_C}\geq\phi k_E+\frac{\partial\hat{b}(P,\theta,\epsilon,X)}{\partial\epsilon}\frac{1}{1+\tau_C},$$

the break-even condition for the banks:

$$\int_\epsilon \hat{b}\left(a,\theta,\epsilon,X_t\right)\mathrm{d}H\left(\epsilon\right)\geq\left(1+(R_F-1)\Delta\right)b\left(a,\theta,X_t\right),$$

and non-negativity constraints on k_E, k_F, k_H, c, c_H, y_E, y_F, ω and P'.

The equilibrium conditions of the model with period length Δ are summarized below.

Definition 4. *Given a sequence of tax rates* $\{\tau_{W,t},\tau_{K,t},\tau_{C,t},\tau_{N,t}\}_{t=0}^{\infty}$*, an* ***equilibrium*** $\mathcal{E}_\Delta$ *of the economy with period length* Δ *is a sequence of prices* $\{R_{F,t},r_{E,t},r_{F,t},w_t\}_{t=0}^{\infty}$*, policy functions giving entrepreneurs' and workers' decisions and a sequence of aggregate variables* $\{C_t,C_{H,t},K_t,$
$K_{E,t},K_{F,t},K_{H,t},Y_t,N_t\}_{t=0}^{\infty}$ *such that:*

1. *The government's budget constraint is balanced every period:*

$$\overline{G}\Delta=\tau_{N,t}w_t\Delta N_t+\tau_{K,t}(Y_t\Delta-w_T\Delta N_t-\delta\Delta K_t)+\tau_{W,t}\Delta K_t+\tau_{C,t}(C_t\Delta-\phi K_{H,t}\Delta).\quad\text{(B.5)}$$

2. *Workers' decision rules solve the worker's optimization problem B.3.*

3. *Entrepreneurs' decision rules are given by the solution to the entrepreneur's problem B.4.*

4. $\{C_t,C_{H,t},K_t,K_{E,t},K_{F,t},K_{H,t},Y_t\}_{t=0}^{\infty}$ *represent the aggregate of household's decisions.*

5. *Newborn agents are indifferent between being entrepreneurs and workers:*

$$\sum_{j=0}^{\infty}\left(1-\rho\Delta\right)^j\left(1-\gamma\Delta\right)^j\log\left(c_{t+j\Delta}^N\right)\Delta=\mathbb{E}_t\left[\sum_{j=0}^{\infty}\left(1-\rho\Delta\right)^j\left(1-\gamma\Delta\right)^j\log\left(c_{i,t+j\Delta}\right)\Delta\right].\quad\text{(B.6)}$$

6. *The asset, intermediate goods and labor markets clear.*

The final goods market clearing condition then follows by Walras' law:

$$\overline{G}_t\Delta+C_t\Delta+K_{t+\Delta}\;=\;Y_t\Delta+\left(1-\delta\Delta\right)\left(K_{E,t}+K_{F,t}-K_{H,t}\right)+C_{H,t}\Delta.\qquad\text{(B.7)}$$

B.2 Solution to the Worker's and Entrepreneur's Problem with Period Length Δ

Following the same derivation as in the main text, where $\Delta = 1$, it can readily be shown that the solution to the worker's problem is given by

$$c^N\Delta = [1-(1-\rho\Delta)(1-\gamma\Delta)]\frac{1+\tilde{R}_F\Delta}{1+\tau_C}P^N = (\gamma+\rho-\gamma\rho\Delta)\frac{1+\tilde{R}_F\Delta}{1+\tau_C}\Delta P^N, \tag{B.8}$$
$$P^{N'} = (1-\rho\Delta)(1+\tilde{R}_F\Delta)P^N. \tag{B.9}$$

Where

$$P^N_{i,t} := a^N_{i,t} + \underbrace{\sum_{j=0}^{\infty}\left[\frac{w_{t+j}\Delta(1-\tau_{N,t+j})(1-\gamma\Delta)^j}{\Pi_{k=0}^{j}1+\tilde{R}_{F,t+k}\Delta}\right]}_{=:F_t}. \tag{B.10}$$

The entrepreneur's between period problem is given by

$$c\Delta = (1-(1-\rho\Delta)(1-\gamma\Delta))\omega = (\gamma+\rho+\gamma\rho\Delta)\Delta\omega, \tag{B.11}$$
$$P' = (1+\tau_C)(1-\rho\Delta)\omega = \frac{1+\tau_C}{1-\gamma\Delta}(\omega - c\Delta). \tag{B.12}$$

The entrepreneur's within period problem is to choose functions $k_E(P,\theta,X)$ and $\omega(P,\theta,X)$ to solve:

$$\sup \int_\epsilon \log(\underline{\omega} + \phi\epsilon k_E)\mathrm{d}H(\epsilon), \tag{B.13}$$

subject to the constraints:

$$\underline{\omega} = \left(-\phi + \frac{1-\tau_K}{1+\tau_C}(r_E\theta - r_F)\Delta\right)k_E + \left(\frac{1}{1+\tau_C}(1+\tilde{R}_F\Delta)\right)P \tag{B.14}$$
$$k_E \geq 0 \tag{B.15}$$
$$\underline{\omega} + \phi\underline{\epsilon}k_E \geq 0, \tag{B.16}$$

where $\tilde{R}_F$ is defined as in the main text. The following proposition summarizes the entrepreneur's optimal decisions.

Proposition 7. *In equilibrium, the entrepreneur's problem has a unique solution for $c(P,\theta,\epsilon,X)$, $P'(P,\theta,\epsilon,X)$, $\omega(P,\theta,\epsilon,X)$ and $k_E(P,\theta,X)$ which depends continuously on the parameters. The entrepreneur's optimal choice of k_E is:*

$$k_E = \frac{S^{-1}\left(\max\left\{0;\min\left\{\frac{(r_E\theta-r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)};S^\star\right\}\right\}\right)P\left(1+\tilde{R}_F\Delta\right)}{\phi - \frac{(r_E\theta-r_F)\Delta(1-\tau_K)}{1+\tau_C}S^{-1}\left(\max\left\{0;\min\left\{\frac{(r_E\theta-r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)};S^\star\right\}\right\}\right)}, \tag{B.17}$$

where $S^\star = S\left(\frac{1}{1-\underline{\epsilon}}\right)$. *For any equilibrium values of* r_E, r_F, *the entrepreneur's choices entail*

$$\omega = \left(\phi\left(\epsilon-1\right)+\frac{1-\tau_K}{1+\tau_C}(r_E\theta-r_F)\Delta\right)k_E+\left(\frac{1}{1+\tau_C}(1+\tilde{R}_F\Delta)\right)P \tag{B.18}$$

$$c = (\gamma+\rho+\gamma\rho\Delta)\omega \tag{B.19}$$

$$P' = (1+\tau_C)(1-\rho\Delta)\omega, \tag{B.20}$$

where

$$S(x)=1-\frac{\int_\epsilon\left(1+x(\epsilon-1)\right)^{-1}\epsilon h\left(\epsilon\right)d\epsilon}{\int_\epsilon\left(1+x(\epsilon-1)\right)^{-1}h\left(\epsilon\right)d\epsilon}, \tag{B.21}$$

and

$$S^\star = S\left(\frac{1}{1-\underline{\epsilon}}\right). \tag{B.22}$$

Here $S:\left[0,\frac{1}{1-\underline{\epsilon}}\right]\to[0,S^\star]$ *is a differentiable and strictly increasing (and therefore invertible) function.*

Proof. The derivative of the entrepreneur's objective function with respect to k_E is:

$$\begin{aligned}
&\frac{\partial}{\partial k_E}\int_\epsilon \log\left(\underline{\omega}+\phi\epsilon k_E\right)\mathrm{d}H(\epsilon)\\
&=\int_\epsilon\left(\underline{\omega}+\phi\epsilon k_E\right)^{-1}\left(\frac{\partial\underline{\omega}}{\partial k_E}+\phi\epsilon\right)\mathrm{d}H(\epsilon)\\
&=\int_\epsilon\left(\underline{\omega}+\phi\epsilon k_E\right)^{-1}\left(\phi(\epsilon-1)+\frac{1-\tau_K}{1+\tau_C}(r_E\theta-r_F)\Delta\right)\mathrm{d}H(\epsilon)\\
&=\left(\underline{\omega}+\phi k_E\right)^{-1}\int_\epsilon\left(\frac{\underline{\omega}}{\underline{\omega}+\phi k_E}+\frac{\phi k_E}{\underline{\omega}+\phi k_E}\epsilon\right)^{-1}\left(\phi(\epsilon-1)+\frac{1-\tau_K}{1+\tau_C}(r_E\theta-r_F)\Delta\right)\mathrm{d}H(\epsilon)\\
&=\left(\underline{\omega}+\phi k_E\right)^{-1}\int_\epsilon(1-x+x\epsilon)^{-1}\left(\phi(\epsilon-1)+\frac{1-\tau_K}{1+\tau_C}(r_E\theta-r_F)\Delta\right)\mathrm{d}H(\epsilon),
\end{aligned}$$

where

$$x=\frac{\phi k_E}{\underline{\omega}+\phi k_E}. \tag{B.23}$$

Here we used that $\underline{\omega}+\phi k_E>0$ at any feasible k_E. This holds because the constraint (B.16) implies that

$$0\le\underline{\omega}+\phi\underline{\epsilon}k_E<\underline{\omega}+\phi\mathbb{E}[\epsilon]k_E=\underline{\omega}+\phi k_E$$

Since $k_E\ge 0$, equation (B.23) in turn implies that $x\ge 0$ at any feasible choice. Furthermore, $x\le\frac{1}{1-\underline{\epsilon}}$. To show this, note that it was shown above that the entrepreneur's end of period consumption is proportional to

$$\omega \equiv \underline{\omega} + \phi\epsilon k_E \equiv (1 - x + x\epsilon)(\underline{\omega} + \phi k_E)$$

Since $\underline{\omega} + \phi k_E > 0$, it follows that $1 - x + x\underline{\epsilon} \geq 0$, or the entrepreneur's consumption would be negative with positive probabilty. Rearranging this condition, we obtain that $x \leq \frac{1}{1-\underline{\epsilon}}$, as desired.

Now we shall show that x is monotonically increasing in k_E. Using the definition of $\underline{\omega}$ in equation (B.16) and the definition of x in equation (B.23), we obtain that

$$\frac{\partial x}{\partial k_E} \propto \left(\frac{1}{1+\tau_C}(1+\tilde{R}_F\Delta)\right)P.$$

Since the entrepreneur can convert units of capital into consumption at rate ϕ and the risk free rate of return is $\frac{1}{1+\tau_C}(1+\tilde{R}_F\Delta)$, there can only be an equilibrium in which some entrepreneurs put a positive amount of capital in the risk-free sector if $\frac{1}{1+\tau_C}(1+\tilde{R}_F\Delta) \geq \phi > 0$. Moreover, by definition $P > 0$, leading us to conclude that $\frac{\partial x}{\partial k_E} > 0$. So x is monotonically increasing in k_E, and $x = 0$ when $k_E = 0$, while $x = \frac{1}{1-\underline{\epsilon}}$ corresponds to the highest possible k_E the entrepreneur can choose while ensuring that consumption is non-negative.

Using that $x \in [0, \frac{1}{1-\underline{\epsilon}}]$, the expression for the derivative of the entrepreneur's objective function can be further rearranged to:

$$\begin{aligned}
&\frac{\phi}{\underline{\omega}+\phi k_E}\int_\epsilon (1+x(\epsilon-1))^{-1}\left(\epsilon - 1 + \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)}\right)\mathrm{d}H(\epsilon) \\
&= \frac{\phi}{\underline{\omega}+\phi k_E}\int_\epsilon (1+x(\epsilon-1))^{-1}\,\mathrm{d}H(\epsilon)\left(\frac{\int_\epsilon (1+x(\epsilon-1))^{-1}\,\epsilon\mathrm{d}H(\epsilon)}{\int_\epsilon (1+x(\epsilon-1))^{-1}\,\mathrm{d}H(\epsilon)} - 1 + \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)}\right) \\
&= \frac{\phi}{\underline{\omega}+\phi k_E}\int_\epsilon (1+x(\epsilon-1))^{-1}\,\mathrm{d}H(\epsilon)\left(\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} - S(x)\right),
\end{aligned}$$

where

$$S(x) = 1 - \frac{\int_\epsilon \Big(1+x(\epsilon-1)\Big)^{-1}\epsilon\cdot\mathrm{d}H(\epsilon)}{\int_\epsilon \Big(1+x(\epsilon-1)\Big)^{-1}\mathrm{d}H(\epsilon)}.$$

Since $\underline{\omega} + \phi k_E > 0$ and $x^{-1} + \epsilon - 1 \geq 0$ for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, with strict inequality for $\epsilon > \underline{\epsilon}$, it follows that the sign of this derivative of the entrepreneur's within period objective function is given by the sign of:

$$\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} - S(x) \tag{B.24}$$

In the case of an interior solution, the first order condition for the optimal choice of k_E is:

$$\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} - S(x) = 0 \tag{B.25}$$

We now show that $S(x)$ is strictly increasing over $x \in (0, \frac{1}{1-\underline{\epsilon}}]$. To this end, we rewrite $S(x)$ as follows:

$$\begin{aligned} S(x) &= 1 - \frac{1}{x}\left(\frac{\int_\epsilon \left(1 + x(\epsilon - 1)\right)^{-1} x\epsilon \cdot \mathrm{d}H(\epsilon)}{\int_\epsilon \left(1 + x(\epsilon - 1)\right)^{-1} \mathrm{d}H(\epsilon)}\right) \\ &= \frac{1}{x} - \left(\int_\epsilon \left(x^{-1} + \epsilon - 1\right)^{-1} \mathrm{d}H(\epsilon)\right)^{-1}. \end{aligned}$$

Since all the terms in $S(x)$ are differentiable with respect to x, for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, it follows that $S(x)$ is itself differentiable over $x \in (0, \frac{1}{1-\underline{\epsilon}}]$. The derivative is:

$$S'(x) = x^{-2}\left(-1 + \left(\int_\epsilon \left(x^{-1} + \epsilon - 1\right)^{-1} \mathrm{d}H(\epsilon)\right)^{-2} \int_\epsilon \left(x^{-1} + \epsilon - 1\right)^{-2} \mathrm{d}H(\epsilon)\right). \tag{B.26}$$

Using Jensen's inequality,

$$\mathbb{E}\left[\left(x^{-1} + \epsilon - 1\right)^{-2}\right] = \mathbb{E}\left[\left(\left(x^{-1} + \epsilon - 1\right)^{-1}\right)^2\right] > \mathbb{E}\left[\left(x^{-1} + \epsilon - 1\right)^{-1}\right]^2.$$

Substituting this into (B.26):

$$S'(x) > x^{-2}\left(-1 + \left(\int_\epsilon \left(x^{-1} + \epsilon - 1\right)^{-1} \mathrm{d}H(\epsilon)\right)^{-2} \left(\int_\epsilon \left(x^{-1} + \epsilon - 1\right)^{-1} \mathrm{d}H(\epsilon)\right)^2\right) = 0,$$

where we used that $x^{-1} + \epsilon - 1 \geq 0$ for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, with strict inequality for $\epsilon > 0$. Therefore, $S(x)$ is strictly increasing over $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, as desired. Equation (B.21) immediately implies that S is continuous over $x \in [0, \frac{1}{1-\underline{\epsilon}}]$ and so $S(x)$ is also strictly increasing over $x \in [0, \frac{1}{1-\underline{\epsilon}}]$.

Recall that the sign of the derivative of the entrepreneur's within-period objective function with respect to k_E is given by (B.24) and that x is monotonically increasing in k_E, with $x = 0$ when $k_E = 0$ and $x = \frac{1}{1-\underline{\epsilon}}$ corresponding to the highest possible k_E the entrepreneur can choose.

Since $S(x)$ is strictly increasing over $x \in [0, \frac{1}{1-\underline{\epsilon}}]$, it follows that there are three cases.

If $\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} \leq S(0)$, then the entrepreneur optimally chooses the corner solution $k_E = x = 0$. If $\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} \geq S\left(\frac{1}{1-\epsilon}\right)$, then the entrepreneur optimally chooses the corner solution $x = \frac{1}{1-\epsilon}$, which corresponds to the highest possible choice of k_E. If $S(0) < \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} < S\left(\frac{1}{1-\epsilon}\right)$, then, by the intermediate value theorem, there is a unique x satisfying the first order condition (B.25). In that case, since the entrepreneur's within-period objective function is strictly concave with respect to k_E, it follows that the first order condition (B.25) characterizes the unique optimal choice of k_E.

For the case where the first order condition holds, we may use the fact that S is monotone and differentiable (and therefore invertible) to rearrange the first order contion as follows:

$$x = S^{-1}\left(\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)}\right)$$

Let $S^\star = S\left(\frac{1}{1-\epsilon}\right)$. Then, $S^{-1}(S^\star) = \frac{1}{1-\epsilon}$. Furthermore, $S(0) = 0$ (by equation (B.21)) and so $S^{-1}(0) = 0$. Since S is monotonically increasing on $x[0, \frac{1}{1-\epsilon}]$, S^{-1} is monotonically increasing on $[0, S^\star]$.

As such, we may group the three cases above as follows:

$$x = \begin{cases} S^{-1}(0) & \text{if } \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} \leq 0 \\ S^{-1}\left(\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)}\right) & \text{if } \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} \in (0, S^\star) \\ S^{-1}(S^\star) & \text{if } \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)} \geq S^\star \end{cases}$$

Combining this with (B.23) to solve for k_E, and simplifying, we have:

$$k_E = \frac{S^{-1}\left(\max\left\{0; \min\left\{\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)}; S^\star\right\}\right\}\right) P\left(1 + \tilde{R}_F\Delta\right)}{\phi - \frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{1+\tau_C} S^{-1}\left(\max\left\{0; \min\left\{\frac{(r_E\theta - r_F)\Delta(1-\tau_K)}{\phi(1+\tau_C)}; S^\star\right\}\right\}\right)}$$

From equations (9) and (B.14), we have that

$$\omega = \left(\phi\left(\epsilon - 1\right) + \frac{1-\tau_K}{1+\tau_C}(r_E\theta - r_F)\Delta\right) k_E + \left(\frac{1}{1+\tau_C}(1 + \tilde{R}_F\Delta)\right) P.$$

Combining these results with equations (B.11) and (B.12) yields all the results of the proposition. □

B.3 Proof of Proposition 1

This proof makes use of the following two lemmas.

Lemma 5. *The following holds, for any* $x \in \left[0; \frac{1}{1-\underline{\epsilon}}\right]$*:*

$$\lim_{\Delta \to 0} \frac{S(x)}{\Delta} = (1-\underline{\epsilon})^2 \varphi^2 x. \tag{B.27}$$

Proof. To prove this, note first that

$$\frac{S(x)}{\Delta} = \frac{1}{\Delta}\left(1 - \frac{\int_{\epsilon} (1+x(\epsilon-1))^{-1} \epsilon \mathrm{d}H(\epsilon)}{\int_{\epsilon} (1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon)}\right) = -\left(\frac{\int_{\epsilon>0} \left(\frac{\epsilon-1}{\Delta}\right)(1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon)}{\int_{\epsilon>0} (1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon)}\right).$$

Then, it remains to show that, for any x in the domain of S,

$$\lim_{\Delta \to 0}\left[\frac{\int_{\epsilon>0} \left(\frac{\epsilon-1}{\Delta}\right)(1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon)}{\int_{\epsilon>0} (1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon)}\right] = -x(1-\underline{\epsilon})^2 \varphi^2. \tag{B.28}$$

To prove this, we prove the following results, from which equation (B.28) follows trivially:

$$\lim_{\Delta \to 0} \int_{\epsilon>0} \left(\frac{\epsilon-1}{\Delta}\right)(1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon) = -x(1-\underline{\epsilon})^2 \varphi^2 \tag{B.29}$$

$$\lim_{\Delta \to 0} \int_{\epsilon>0} (1+x(\epsilon-1))^{-1} \mathrm{d}H(\epsilon) = 1 \tag{B.30}$$

To prove these, recall that the $\epsilon_{i,t}$ is given by:

$$\epsilon_{i,t} = \underline{\epsilon} + (1-\underline{\epsilon}) \exp\left(\varphi \sqrt{\Delta} \mathcal{R}_{i,t} - \frac{\varphi^2 \Delta}{2}\right).$$

where $\mathcal{R}_{i,t} \sim N(0,1)$. This implies that

$$\epsilon_{i,t} - 1 = (1-\underline{\epsilon})\left(\exp\left(\varphi \sqrt{\Delta} \mathcal{R}_{i,t} - \frac{\varphi^2 \Delta}{2}\right) - 1\right).$$

Therefore, the left hand side of (B.29) is:

$$\lim_{\Delta \to 0} \int_{-\infty}^{\infty} (1-\underline{\epsilon}) \left(\frac{\exp\left(\varphi\sqrt{\Delta}\mathcal{R} - \frac{\varphi^2\Delta}{2}\right) - 1}{\Delta}\right)\left(1 + x(1-\underline{\epsilon})\left(\exp\left(\varphi\sqrt{\Delta}\mathcal{R} - \frac{\varphi^2\Delta}{2}\right) - 1\right)\right)^{-1} \mathrm{d}\Phi(\mathcal{R}),$$

where $\Phi(\cdot)$ is the standard normal cdf. Likewise, the left hand side of (B.30) is:

$$\lim_{\Delta \to 0} \int_{-\infty}^{\infty} \left(1 + x(1-\underline{\epsilon})\left(\exp\left(\varphi\sqrt{\Delta}\mathcal{R} - \frac{\varphi^2\Delta}{2}\right) - 1\right)\right)^{-1} \mathrm{d}\Phi(\mathcal{R}).$$

Now, we consider a first order approximation of $\exp\left(\varphi\sqrt{\Delta}\mathcal{R} - \frac{\varphi^2\Delta}{2}\right)$ in units of $\sqrt{\Delta}$, around the point $\sqrt{\Delta} = 0$:

$$\exp\left(\varphi\sqrt{\Delta}\mathcal{R} - \frac{\varphi^2\Delta}{2}\right) \simeq 1 + \varphi\sqrt{\Delta}\mathcal{R}. \tag{B.31}$$

Similarly, in the neighborhood of $\sqrt{\Delta} = 0$:

$$\left(1 + x(1-\underline{\epsilon})\left(\exp\left(\varphi\sqrt{\Delta}\mathcal{R} - \frac{\varphi^2\Delta}{2}\right) - 1\right)\right)^{-1} \simeq 1 - x(1-\underline{\epsilon})\varphi\sqrt{\Delta}\mathcal{R}. \tag{B.32}$$

Multiplying the term in the limit on the left hand side of (B.29) by Δ, and ignoring terms of order greater than Δ, we therefore can write it as:

$$\begin{aligned}
&\int_{-\infty}^{\infty} (1-\underline{\epsilon})\varphi\sqrt{\Delta}\cdot\mathcal{R}\left(1 - x(1-\underline{\epsilon})\varphi\sqrt{\Delta}\mathcal{R}\right) \mathrm{d}\Phi\left(\mathcal{R}\right) \\
&= \int_{-\infty}^{\infty} (1-\underline{\epsilon})\varphi\sqrt{\Delta}\cdot\mathcal{R} - x(1-\underline{\epsilon})^2\varphi^2\Delta\mathcal{R}^2 \mathrm{d}\Phi\left(\mathcal{R}\right) \\
&= (1-\underline{\epsilon})\varphi\sqrt{\Delta}E[\mathcal{R}] - x(1-\underline{\epsilon})^2\varphi^2\Delta E[\mathcal{R}^2] \\
&= -x(1-\underline{\epsilon})^2\varphi^2\Delta.
\end{aligned}$$

since $\mathcal{R}$ is normally distributed. Hence (B.29) follows immediately.

Likewise, considering the term on the left hand side of (B.30), we have:

$$\begin{aligned}
&\int_{-\infty}^{\infty} 1 - x(1-\underline{\epsilon})\varphi\sqrt{\Delta}\mathcal{R}\mathrm{d}\Phi\left(\mathcal{R}\right) \\
=&1 - x(1-\underline{\epsilon})\varphi\sqrt{\Delta}E[\mathcal{R}] \\
=&1.
\end{aligned}$$

This proves (B.30). □

Lemma 6. *For any $z \in (-\infty, \infty)$, it holds that*

$$\lim_{\Delta\to 0} S^{-1}(\max\{0; \min\{\Delta z; S^\star\}\}) = \max\left\{0; \min\left\{\frac{z}{(1-\underline{\epsilon})^2\varphi^2}; \frac{1}{1-\underline{\epsilon}}\right\}\right\}. \tag{B.33}$$

Proof. To prove (B.33), we first note that

$$\lim_{\Delta\to 0} \frac{S^\star}{\Delta} = (1-\underline{\epsilon})\varphi^2. \tag{B.34}$$

This follows immediately from the definition of $S^\star$ in (B.22) and from Lemma 5.

Now, we show that

$$\forall z \in [0,(1-\underline{\epsilon})\varphi^2), \ \lim_{\Delta\to 0} S^{-1}(\Delta z) = \frac{z}{\varphi^2(1-\underline{\epsilon})^2}. \tag{B.35}$$

To prove (B.35) , define the function $\mathscr{F}(x)$ according to:

$$\mathscr{F}(x) = \frac{S(x)}{\Delta}. \tag{B.36}$$

where $x \in \left(0, \frac{1}{1-\underline{\epsilon}}\right)$. Given that S is continuous and strictly increasing, it follows that function $\mathscr{F}(\cdot)$ is continuous and strictly increasing, and therefore invertible. We now show that, for any z in the range of $\mathscr{F}$,

$$\mathscr{F}^{-1}(z) \equiv S^{-1}(\Delta z). \tag{B.37}$$

To show this let $x = \mathscr{F}^{-1}(z)$. Then $z = \mathscr{F}(x) = \frac{S(x)}{\Delta}$, so $S(x) = \Delta z$ and $x = S^{-1}(\Delta z)$, giving us (B.37).

Define the function $\overline{\mathscr{F}}(x) = (1-\underline{\epsilon})^2\varphi^2 x$. It follows from simple rearrangement that its inverse is:

$$\overline{\mathscr{F}}^{-1}(x) = \frac{x}{(1-\underline{\epsilon})^2\varphi^2}. \tag{B.38}$$

We know from Lemma 5 that, as $\Delta \to 0$, $\mathscr{F}(x)$ converges to $\overline{\mathscr{F}}(x)$. Since $\mathscr{F}$ is continuous, this convergence is uniform, and the inverse $\mathscr{F}^{-1}(x)$ converges to $\overline{\mathscr{F}}^{-1}(x)$. Then, using equations (B.37) and (B.38), we obtain (B.35) for values of z in the relevant domain.

Note that, for $\Delta > 0$, the domain of $S^{-1}(\cdot)$ is $[0, S^\star]$. Therefore, the result (B.35) must hold for all $z \in \left(0, \lim_{\Delta\to 0}\frac{S^\star}{\Delta}\right) \equiv (0,(1-\underline{\epsilon})\varphi^2)$ since, for any such z, Δz will be in will be in the domain of S^{-1} for sufficiently small $\Delta > 0$. Equally, it must be true that $\lim_{\Delta\to 0} S^{-1}(0) = 0$, since $S^{-1}(0) = 0$ for any value of $\Delta > 0$. Therefore, (B.35) follows for $z \in [0,(1-\underline{\epsilon})\varphi^2)$.

Now, using (B.34) and (B.35), we prove (B.33) for $z \in (-\infty, \infty)$. We proceed in cases. First, consider the case $z \leq 0$. Then, for sufficiently small $\Delta > 0$, equation (B.34) implies that $\Delta z \leq 0 < S^\star$. Then,

$$\lim_{\Delta\to 0} S^{-1}(\max\{0;\min\{\Delta z; S^\star\}\}) = \lim_{\Delta\to 0} S^{-1}(0) = 0, \tag{B.39}$$

where the second equality used (B.35).

Now, suppose that $z \in \left(0, \frac{1}{1-\underline{\epsilon}}\right)$. Then, for sufficiently small $\Delta > 0$, equation (B.34) implies that $0 < \Delta z < S^\star$. Then,

$$\lim_{\Delta\to 0} S^{-1}(\max\{0;\min\{\Delta z; S^\star\}\}) = \lim_{\Delta\to 0} S^{-1}(\Delta z) = \frac{z}{(1-\underline{\epsilon})^2\varphi^2}, \tag{B.40}$$

where the second equality used (B.35).

Now, suppose that $z > \frac{1}{1-\underline{\epsilon}}$. Then, for sufficiently small $\Delta > 0$, equation (B.34) implies that $\Delta z > S^\star > 0$. Then,

$$\lim_{\Delta\to 0} S^{-1}(\max\{0; \min\{\Delta z; S^\star\}\}) = \lim_{\Delta\to 0} S^{-1}(S^\star) = \lim_{\Delta\to 0} \frac{1}{1-\underline{\epsilon}} = \frac{1}{1-\underline{\epsilon}}, \tag{B.41}$$

where the second equality used that, for any $\Delta > 0$, $S^{-1}(S^\star) = \frac{1}{1-\underline{\epsilon}}$, since $S^\star = S\left(\frac{1}{1-\underline{\epsilon}}\right)$, by definition.

Comparing equation (B.33) with equations (B.39),(B.40) and (B.41), we see that we have proven (B.33) for any $z \in (-\infty, \infty)$ except for $z = \frac{1}{1-\underline{\epsilon}}$. In particular, (B.33) must hold for all $z \neq \frac{1}{1-\underline{\epsilon}}$ in the neighborhood of $\frac{1}{1-\underline{\epsilon}}$. Now, note that the left hand side of (B.33) is continuous and weakly increasing in z for any $\Delta > 0$. Equally, the right hand side of (B.33) is continuous and weakly increasing in z. Continuity arguments then imply that (B.33) also holds at $z = \frac{1}{1-\underline{\epsilon}}$. □

Proof of the Propositon

First, derive the worker's continuous time solution. Consider the discrete time solution with period length Δ.

$$c^N \Delta = (\gamma + \rho - \gamma\rho\Delta) \frac{1 + \tilde{R}_F \Delta}{1 + \tau_C} \Delta P^N, \tag{B.42}$$

$$P^{N'} = (1 - \rho\Delta)(1 + \tilde{R}_F \Delta) P^N, \tag{B.43}$$

Taking the limit of equation (B.42) as $\Delta \to 0$, we obtain equation (12).

Now consider

$$\frac{P^{N'} - P^N}{\Delta} = \left[(1 - \rho\Delta)\tilde{R}_F - \rho\right] P^N.$$

Taking the limit of this as $\Delta \to 0$, we get

$$dP^N = \left\{\left[\tilde{R}_F - \rho\right] P^N\right\} \mathrm{d}t.$$

Using the solution for c^N, given by equation (12), we obtain (13).

Now derive the entrepreneur's continuous time solution. Taking the limit of equation (B.17) as $\Delta \to 0$, and using Lemma 6, we obtain (14). Combining (B.18) and (B.19), we get:

$$c = (\rho + \gamma - \rho\gamma\Delta) \left[P \left(\frac{1 + \tilde{R}_F \Delta}{1 + \tau_C}\right) + k_E \left(\phi(\epsilon - 1) + \frac{1 - \tau_K}{1 + \tau_C}(r_E\theta - r_F)\Delta\right)\right].$$

Taking the limit of this as $\Delta \to 0$ and noting that, as $\Delta \to 0$, $\epsilon \to 1$ in probability, we obtain

(15). Finally, we note that Proposition 7 implies that:

$$P' = \frac{1+\tau_C}{1-\gamma\Delta}\left(\left[P\left(\frac{1+\tilde{R}_F\Delta}{1+\tau_C}\right)+k_E\left(\phi\left(\epsilon-1\right)+\frac{1-\tau_K}{1+\tau_C}(r_E\theta-r_F)\Delta\right)\right]-c\Delta\right),$$

so that

$$\frac{P'-P}{\Delta} = \frac{1+\tau_C}{1-\gamma\Delta}\left(P\left(\frac{\tilde{R}_F+\gamma}{1+\tau_C}\right)+k_E\left[\phi\left(\frac{\epsilon-1}{\Delta}\right)+\frac{1-\tau_K}{1+\tau_C}(r_E\theta-r_F)\right]-c\right). \quad \text{(B.44)}$$

Now, for sufficiently small Δ, $\frac{\epsilon i,t-1}{\Delta} \simeq (1-\underline{\epsilon})\varphi\sqrt{\Delta}\mathcal{R}_{i,t}$. Furthermore, $\frac{\mathcal{R}_{i,t}}{\sqrt{\Delta}}$ corresponds to the difference of a standard Brownian motion, since $\frac{\mathcal{R}_{i,t}}{\sqrt{\Delta}} \sim N(0,\Delta)$. Therefore, as $\Delta \to 0$, equation (B.44) simplifies to (16). □

B.4 Proof of Proposition 2

As a first step, we characterize a steady state equilibrium in Proposition 8, as per Definitions 1 and 2. As a second step, we show that the aggregate variables $\{K^*, K_E^*, C^*, N^*, \mathcal{R}^*, Y^*\}$, post-tax prices $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ and consumption tax rate τ_C^* constitute an equilibrium according to Proposition 2 if and only if the aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_C^*, \tau_N^*\}$ constitute an equilibrium according to Proposition 8.

B.4.1 Alternative characterization of the steady state equilibrium

Proposition 8. *There exists a steady state $\mathscr{S}$ which is consistent with the particular values of aggregate variables aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_C^*, \tau_N^*\}$ and in which no entrepreneurs hide capital or intermediate goods, if and only if the following conditions hold:*

1. *The government's budget constraint is balanced every period:*

$$\bar{G} = \tau_N^* w^* N^* + \tau_K^*(\pi_F^*(1-N^*) + (r_E^*-r_F^*)K_E^* + (r_F^*-\delta)K^*) + \tau_W^* K^* + \tau_C^* C^* \quad \text{(B.45)}$$

This comes directly from the governement budget constraint of an economy with period length Δ given by equation (B.5).

2. *Aggregate consumption respects the solution to the worker's and entrepreneur's optimization problem:*

$$C^* = \frac{\rho+\gamma}{1+\tau_C^*}\left(K^* + N^* F^{N*} + (1-N^*)F^*\right), \quad \text{(B.46)}$$

We derive aggregate steady state consumption (B.46) by integrationg over the consumption policy function of workers and entrepreneurs, given by equations (12) and (15):

$$C^* = \frac{\rho+\gamma}{1+\tau_C^*}(\int_{i\leq N^*} P_i^{N*} di + \int_{i>N^*} P_i^* di),$$
$$C^* = \frac{\rho+\gamma}{1+\tau_C^*}(\int_{i\leq N^*} (a_i^{N*} + F^{N*}) di + \int_{i>N^*} (a_i^* + F^*) di),$$
$$C^* = \frac{\rho+\gamma}{1+\tau_C^*}(K^* + N^* F^{N*} + (1-N^*)F^*).$$

Where F^{N*} and F^* are obtained by evaluating (B.2) and (B.10) at steady state prices and steady state tax rates and taking the limit as $\Delta \to 0$:

$$F^{N*} = \frac{w^*(1-\tau_N^*)}{\gamma + (R_F^* - 1)} \tag{B.47}$$
$$F^* = \frac{\pi_F^*(1-\tau_K^*)}{\gamma + (R_F^* - 1)} \tag{B.48}$$

3. *Aggregate risky capital respects the solution to the entrepreneur's optimization problem:*

$$K_E^* = \int_{i>N^*} \hat{k}_E(\theta_i) P_i^* di = \sum_\theta \hat{k}_E(\theta) \int_{i>N^*,\theta_i=\theta} P_i^* di = \mathbb{P}^* \sum_\theta \hat{k}_E(\theta)\mu^*(\theta)$$
$$= \mathbb{P}^* \hat{k}_E(1)\mu^*(1), \tag{B.49}$$

where $\hat{k}_E(1)$ is given by the continuous time policy function (14), and we directly use that optimally $\hat{k}_E(0) = 0$.

4. *No-Arbitrage condition: entrepreneurs are indifferent between investing in the production of the risk-free intermediate good and lending capital to the bank:*

$$R_F^* = 1 + \left[(1-\tau_K^*)(r_F^* - \delta) - \tau_W^*\right]. \tag{B.50}$$

5. *The stationary distribution of entrepreneurial wealth respects transition probabilites and entrepreneur's policy functions:*

   (a) *Total entrepreneurial wealth, $\mathbb{P}^*$, is given by:*

$$\mathbb{P}^* = \frac{1}{\gamma + \rho + (R_F^* - 1) - (r_E^* - r_F^*)(1-\tau_K^*)\hat{k}_E(1)\mu^*(1)} \gamma(1-N^*)F^* \tag{B.51}$$

(b) *The fraction of entrepreneurial wealth owned by entrepreneurs of type θ, $\mu^*(\theta)$, is given by:*

$$\mu^*(1) = \frac{\lambda_\theta g(1) + \gamma g(1)\frac{(1-N^*)F^*}{\mathbb{P}^*}}{\gamma + \rho + \lambda_\theta - \left((R_F^* - 1) + \hat{k}_E(1)\left(r_E^* - r_F^*\right)\left(1 - \tau_K^*\right)\right)} \tag{B.52}$$

$$\mu^*(0) = 1 - \mu^*(1)$$

The details of the derivation of the above characterizations of steady state wealth can be found in Appendix B.4.4.

6. *Optimality in the production of the final good.*[17] *Moreover the intermediate goods markets and the labor market clear:*

$$\pi_F^* = f\left(\frac{K_E^*}{1-N^*}, \frac{K^*-K_E^*}{1-N^*}, \frac{N^*}{1-N^*}\right) - r_E^*\frac{K_E^*}{1-N^*} - r_F^*\frac{K^*-K_E^*}{1-N^*} - w^*\frac{N^*}{1-N^*}, \tag{B.53}$$

$$r_E^* = f_1\left(\frac{K_E^*}{1-N^*}, \frac{K^*-K_E^*}{1-N^*}, \frac{N^*}{1-N^*}\right), \tag{B.54}$$

$$r_F^* = f_2\left(\frac{K_E^*}{1-N^*}, \frac{K^*-K_E^*}{1-N^*}, \frac{N^*}{1-N^*}\right), \tag{B.55}$$

$$w^* = f_3\left(\frac{K_E^*}{1-N^*}, \frac{K^*-K_E^*}{1-N^*}, \frac{N^*}{1-N^*}\right). \tag{B.56}$$

7. *Newborn agents are indifferent between being entrepreneurs and workers:*

$$\log(\tilde{w}^*) = \log(\tilde{\pi}_F^*) + \frac{g(1)}{\rho+\gamma}\left[(r_E^* - r_F^*)(1-\tau_K^*)\hat{k}(1) - \frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(1)\varphi\right)^2\right]. \tag{B.57}$$

This condition follows from substituting the steady state expressions for $V^N\left(F^{N*}, X^*\right)$ and $V\left(F^*, \theta, X^*\right)$ (derived in Appendix B.4.5 and B.4.6) into the indifference condition of a newborn:

$$V^N\left(F^{N*}, X^*\right) = \mathbb{E}_\theta V\left(F^*, \theta, X^*\right).$$

8. *The final goods market clears:*

$$C^* + \delta K^* + \overline{G} = f\left(\frac{K_E^*}{1-N^*}, \frac{K^*-K_E^*}{1-N^*}, \frac{N^*}{1-N^*}\right)(1-N^*), \tag{B.58}$$

which is obtained by evaluating the goods market clearing condition at steady state

[17] Note that this leads to equalized output f and profits π_F^* from final good production across all entrepreneurs.

values.

9. *Lastly, the following inequality conditions are satisfied:*

$$r_E^* > r_F^*, \quad K_E^* < K^*, \quad \tau_C^* > -1, \tag{B.59}$$

$$\phi \leq \frac{1-\tau_K^*}{1+\tau_C^*} \text{ with strict inequality if } r_E^* < \delta. \tag{B.60}$$

Note that the asset market clearing condition holds by Walras' law.

B.4.2 Showing equivalence

Proposition 9. *The aggregate variables* $\{K^*, K_E^*, C^*, N^*, \mathcal{R}^*, Y^*\}$*, post-tax prices* $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ *and consumption tax rate* τ_C^* *constitute a steady state according to Proposition 2 if and only if the aggregate variables* $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$*, prices* $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ *and taxes* $\{\tau_W^*, \tau_K^*, \tau_C^*, \tau_N^*\}$ *constitute a steady state according to Proposition 8.*

Proof. $\Leftarrow$: Suppose that aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_C^*, \tau_N^*\}$ satisfy the equilibrium conditions in Proposition 8, (B.45) - (B.58), and the inequalities (B.59) - (B.60). Define post-tax prices as in equations (17)-(20): $\tilde{w} = \frac{w(1-\tau_N)}{1+\tau_C}$, $\tilde{\pi}_F = \frac{(1-\tau_K)\pi_F}{1+\tau_C}$, $\tilde{R}_F = R_F - 1$ and $\tilde{r}_X = \frac{(r_E - r_F)(1-\tau_K)}{1+\mathcal{R}(K_E, K_F, N, \pi_0)}$, where $\mathcal{R}(\cdot)$ satisfies equation (21).

Replace any occurance of pre-tax prices in the equilibrium conditions with their post-tax price counterpart, noting that we may replace $\{r_E^*, r_F^*\}$ with $\tilde{r}_X^*$ because the two prices only appear as differences. The new set of inequality conditions follows directly. By eliminating the variables $\left\{F^*, F^{N*}, \mathbb{P}^*\right\}$ and rearranging, we may arrive at the set of equilibrium conditions of Proposition 2, (22)-(27).

$\Rightarrow$: Suppose that aggregate variables $\{K^*, K_E^*, C^*, N^*, \mathcal{R}^*, Y^*\}$, pre-tax prices $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ and consumption tax τ_C^* satisfy the equilibrium conditions of Proposition 2, (22)-(27). Then there exists an appropriate choice for $\left\{F^*, F^{N*}, \mathbb{P}^*\right\}$, prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_N^*\}$ such that the full set of equilibrium variables satisfies the conditions of Proposition 8. □

B.4.3 Derivation of the inequalities in Proposition 8 (equations B.59 - B.60)

We claim that the following conditions are necessary for an equilibrium: $r_E^* > r_F^*$, $K_E^* < K^*$, $\tau_C^* > -1$ and $\phi \leq \frac{1-\tau_K^*}{1+\tau_C^*}$ with strict inequality if $r_E^* < \delta$.

1. $r_E^* > r_F^*$: By contradiction, suppose $r_E^* \leq r_F^*$. Then entrepreneurs will strictly prefer to set $k_E = 0$. By the Inada conditions of the final good production function, r_E^* goes to infinity.

2. $K_E^* < K^*$: Otherwise $K_F^* \leq 0$, and so $r_F^*(1-\tau_K^*)$ is either infinity or undefined.

3. $\tau_C^* > -1$: Otherwise, consumption will be unbounded, violating feasibility.

4. $\phi \leq \frac{1-\tau_K^*}{1+\tau_C^*}$ with strict inequality if $r_E^*(1-\tau_K^*) < \delta\frac{1-\tau_K^*}{1+\tau_C^*}$:
   - The equivalent of this condition for an economy with period length Δ is $r_E^*(1-\tau_K^*)\Delta + (1-\delta\Delta)\frac{1-\tau_K^*}{1+\tau_C^*} \geq \phi$.
   - This condition is necessary to rule out the hiding of capital by entrepreneurs investing in the risky technology.[18]
   - Taking the limit as $\Delta \to 0$, we obtain the required result.

B.4.4 Derivation of $\mathbb{P}^*$ and $\mu^*(\theta)$ (equations B.51 andb B.52)

Derivation of $\mathbb{P}^*$. Note that every period a fraction $\gamma\Delta$ of entrepreneurs die and a fraction $(1-N)$ of the $\gamma\Delta$ newborn agents choose to become entrepreneurs, each with initial wealth F. This gives rise to the following law of motion for $\mathbb{P}$:

$$\mathbb{P}' = (1-\gamma\Delta)\int_{i>N} P_i di + \gamma\Delta(1-N)F \tag{B.61}$$

$$\mathbb{P}' = (1-\gamma\Delta)(1-\rho\Delta)\left[(r_E^* - r_F^*)\Delta(1-\tau_K^*)\hat{k}_E(1)\mu^* + (1+\tilde{R}_F^*\Delta)\right]\mathbb{P} + \gamma\Delta(1-N)F,$$

Note that we evaluate the integral $\int_{i>N} P_i di$ by combining the expression ω (equation B.18) and the entrepreneur's policy function for $\mathbb{P}^*$ (equation B.20) and the entrepreneur's policy function for k_E (equation B.17).

Using that in steady state $\mathbb{P}' = \mathbb{P} = \mathbb{P}^*$ and taking the limit as $\Delta \to 0$, we obtain equation B.51, as required:

$$\mathbb{P}^* = \frac{1}{\gamma + \rho + \tilde{R}_F^* - (r_E^* - r_F^*)(1-\tau_K^*)\hat{k}_E(1)\mu^*(1)}\gamma(1-N^*)F^*.$$

Derivation of $\mu^*(\theta)$. Recall that we define $\mu^*(\theta)$ as the fraction of total entrepreneur wealth, $\mathbb{P}^*$, held by entrepreneurs of type θ in the steady state.

$$\mu^*(\theta) := \frac{\int_{i>N^*,\theta_i=\theta} P_i^* di}{\mathbb{P}^*} \tag{B.62}$$

[18] To be precise, this condition only applies when the entrepreneur does not engage with the bank. Whenever the entrepreneur borrows from or lends to the bank, they will agree on an incentive compatible contract.

The law of motion for $\mu'(\theta)$ iin an economy of length Δ is given by:

$$\begin{aligned}\mu'(\theta)\mathbb{P}' =& \lambda_\theta \Delta g(\theta)(\mathbb{P}' - \gamma\Delta(1-N)F) + \gamma\Delta g(\theta)(1-N)F \\ &+ (1-\gamma\Delta)(1-\lambda_\theta\Delta)\mu(\theta)\mathbb{P}\int_\epsilon \hat{P}'(\theta,\epsilon)\mathrm{d}H(\epsilon)\end{aligned} \tag{B.63}$$

At the end of each period, entrepreneurs die with probability $\gamma\Delta$. Out of the $(1-\gamma\Delta)$ entrepreneurs that survive, they keep the same θ with probability $(1-\lambda_\theta\Delta)$ and draw a new θ with probability $\lambda_\theta\Delta$. The left hand side corresponds to the total capital held by entrepreneurs of type θ in the next period. The three right hand terms are derived as follows:

1. The first term indicates the capital held in the next period by surviving entrepeneurs who draw a new ability level that period and happen to draw θ. In total, surviving entrepreneurs hold $\mathbb{P}' - \gamma\Delta(1-N)F$ of capital. Since the probability of drawing a new ability level is orthogonal to ability type, the fraction of entrepreneurs drawing theta as their new type is equal to the fraction of wealth these entrepreneurs own.

2. The second term is given by the wealth owned by newborns, $(1-N)F$, scaled by the fraction of newborns who happen to draw θ.

3. Fraction $(1-\gamma\Delta)(1-\lambda_\theta\Delta)$ of entrepreneurs of type θ surive and retain their type in the next period. Due to the linearity of the policy function in P, the cumulative wealth of these entrepreneurs in the next period is equal to the expected wealth of an entrepreneur who owns their cumulative wealth, $(1-\gamma\Delta)(1-\lambda_\theta\Delta)\mu(\theta)\mathbb{P}$. Note that we have defined $\hat{P}'(\theta,\epsilon) := \frac{P'}{P}$.

In a steady state, $\mu' = \mu = \mu^*$ and $\mathbb{P}' = \mathbb{P} = \mathbb{P}^*$. In that case, equation (B.63) simplifies to:

$$\mu^*(\theta) = \frac{\lambda_\theta\Delta g(\theta) + \gamma\Delta(1-N^*)\frac{F^*}{\mathbb{P}^*}g(\theta)(1-\lambda_\theta\Delta)}{1-(1-\gamma\Delta)(1-\lambda_\theta\Delta)\int_\epsilon \hat{P}'(\theta,\epsilon)\mathrm{d}H(\epsilon)} \tag{B.64}$$

To find the expression for $\int_\epsilon \hat{P}'(\theta,\epsilon)\mathrm{d}H(\epsilon)$, we plug the expression for ω given by equation (B.18) into the entrepreneur's policy function (B.20). We then use the fact that k_E is linear in P and that $\mathbb{E}[\epsilon] = 1$.

$$\begin{aligned}\int_\epsilon \hat{P}'(\theta,\epsilon)\mathrm{d}H(\epsilon) &= (1+\tau_C)(1-\rho\Delta)\int_\epsilon \omega(\theta,\epsilon,1)\mathrm{d}H(\epsilon) \\ &= (1+\tau_C)(1-\rho\Delta)\int_\epsilon\Big((\phi(\epsilon-1) + \frac{1-\tau_K}{1+\tau_C}(r_E\theta - r_F)\Delta)\hat{k}_E(\theta) + \frac{1+\tilde{R}_F\Delta}{1+\tau_C}\Big)\mathrm{d}H(\epsilon) \\ &= (1-\rho\Delta)\Big((r_E\theta - r_F)(1-\tau_K)\Delta\hat{k}_E(\theta) + (1+\tilde{R}_F\Delta)\Big)\end{aligned}$$

Taking the limit as $\Delta \to 0$, we obtain equation (B.52), as required:

$$\mu^*(\theta) = \frac{\lambda_\theta g(\theta) + \gamma g(\theta) \frac{(1-N^*)F^*}{\mathbb{P}^*}}{\gamma + \rho + \lambda_\theta - \Big((R_F^* - 1) + \hat{k}_E(1)\left(r_E^* - r_F^*\right)(1 - \tau_K^*) \Big)}. \tag{B.65}$$

B.4.5 Derivation of the worker's steady state value function

In this section, we show that that the worker's steady state value function is given by:

$$(\gamma + \rho) V^N\left(F^{N*}, X^*\right) = \log\left(\frac{\gamma + \rho}{1 + \tau_C} F^N\right) + \frac{(1 + \tau_C)(\tilde{r}_F - \tilde{p}) - \rho}{\gamma + \rho} \tag{B.66}$$

We shall this continuous time Bellman equation by taking the limit of the Bellman equation of the problem for a given period length Δ:

$$V^N(P, X^*) = \max\Big(\log(c^N)\Delta + (1 - \rho\Delta)(1 - \gamma\Delta) V^N\left(P^{N\prime}, X^*\right) \Big)$$
$$(\rho + \gamma - \rho\gamma\Delta) V^N(P, X^*) = \max\Big(\log(c^N) + (1 - \rho\Delta)(1 - \gamma\Delta) \frac{V^N\left(P^{N\prime}, X^*\right) - V^N(P, X^\star)}{\Delta} \Big)$$

Hence, taking the limit as $\Delta \to 0$, and using equation (12) to substitute in for the continuous time policy function for c^N, we obtain

$$(\rho + \gamma) V^N(P, X^*) = \log\left(\frac{\rho + \gamma}{1 + \tau_C} P^N\right) + \underbrace{\lim_{\Delta \to 0} \left(\frac{V^N\left(P^{N\prime *}, X^*\right) - V^N(P, X^*)}{\Delta} \right)}_{\text{A}} \tag{B.67}$$

As the next step, we compute the term A. As in Appendix A.1, we can write the worker's steady state value function in an economy with period length Δ as:

$$V^N(P, X^*) = Q(\Delta) + \frac{1}{1 - (1 - \rho\Delta)(1 - \gamma\Delta)} \log(P^N)\Delta. \tag{B.68}$$

Using the policy function for P' for the economy of period length Δ, given by equation (B.9), and l'Hôpital's rule we obtain:

$$\begin{aligned} A &= \lim_{\Delta \to 0} \left(\frac{1}{1 - (1 - \rho\Delta)(1 - \gamma\Delta)} \log\Big((1 - \rho\Delta)(1 + (R_F - 1)\Delta) \Big) \right) \\ &= \lim_{\Delta \to 0} \left(\frac{1}{\rho + \gamma - 2\rho\gamma\Delta} \frac{-\rho(1 + (R_F - 1)\Delta) + (1 - \rho\Delta)(R_F - 1)}{(1 - \rho\Delta)(1 + (R_F - 1)\Delta)} \right) \\ &= \frac{R_F - 1 - \rho}{\rho + \gamma}. \end{aligned}$$

B.4.6 Derivation of the entrepreneur's steady state value function

In this section we show that the entrepreneur's steady state value function is given by:

$$
\begin{aligned}
(\gamma+\rho)V(F,\theta,X^*) &= \log\left(\frac{\gamma+\rho}{1+\tau_C}F\right)+\frac{\tilde{R}_F-\rho}{\gamma+\rho} \\
&+\frac{1}{\gamma+\rho}\frac{\gamma+\rho+\lambda_\theta g(\theta)}{\gamma+\rho+\lambda_\theta}\left((r_E\theta-r_F)(1-\tau_K)\hat{k}(\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(\theta)\varphi\right)^2\right) \\
&+\frac{1}{\gamma+\rho}\frac{\lambda_\theta g(1-\theta)}{\gamma+\rho+\lambda_\theta}\left((r_E(1-\theta)-r_F)(1-\tau_K)\hat{k}(1-\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(1-\theta)\varphi\right)^2\right).
\end{aligned}
\tag{B.69}
$$

We shall derive this continuous time Bellman equation by taking the limit of the Bellman equation of the problem for a given period length Δ. We evaluate equation B.4 at $X = X' = X^*$ and substitute in the policy functions for c and P':

$$
\begin{aligned}
V(P,\theta,X^*) &= \int_{\epsilon>0}\Bigg(\log(c(P,\theta,\epsilon,X^*)))\Delta+(1-\rho\Delta)(1-\gamma\Delta) \\
&\times\ E\left[V(P'(P,\theta,\epsilon,X^*)),\theta',X^*)\middle|\theta\right]\Bigg)\mathrm{d}H(\epsilon)
\end{aligned}
\tag{B.70}
$$

Now using the process for θ, we may write out the conditional expectation as

$$
\begin{aligned}
E\left[V(P',\theta',X^*)\middle|\theta\right] =&[1-\lambda_\theta\Delta(1-g(\theta))]V(P',\theta,X^*) \\
&+\lambda_\theta\Delta(1-g(\theta))V(P',(1-\theta),X^*)
\end{aligned}
$$

As in Lemma 1, we may write the entrepreneur's value function in an economy with period length Δ as:

$$
V(P,\theta,X)=\bar{V}(\theta,X,\Delta)+\frac{1}{1-(1-\rho\Delta)(1-\gamma\Delta)}\log(P)\Delta. \tag{B.71}
$$

Hence, as $\Delta\to 0$, we obtain the continuous time entrepreneur value function:

$$
V(P,\theta,X^*)=\lim_{\Delta\to 0}\bar{V}(\theta,X^*,\Delta)+\frac{1}{\rho+\gamma}\log(P). \tag{B.72}
$$

What is left to do is to derive $\bar{V}(\theta,X,\Delta)$ and take the limit as $\Delta\to 0$. Since we have already derived the solutions for c, P' and k_E, deriving $\bar{V}(\theta,X,\Delta)$ is simply a matter of substituting these expressions into the value function and rearranging. To simplify the algebra, define $\alpha_\Delta := (1-\rho\Delta)(1-\gamma\Delta)$.

$$\omega = \underline{\omega} + \theta\epsilon k_E \underset{(B.18),(B.17)}{=} \underbrace{\left((\phi(\epsilon-1) + \frac{1-\tau_K}{1+\tau_C}(r_E\theta - r_F)\Delta)\hat{k}_E(\theta) + \frac{1+\tilde{R}_F\Delta}{1+\tau_C}\right)}_{=:D(\theta,\epsilon,\Delta)} P \quad \text{(B.73)}$$

$$c \underset{(B.11)}{=} (\rho + \gamma + \rho\gamma\Delta)\omega = (\rho + \gamma + \rho\gamma\Delta)D(\theta,\epsilon,\Delta)P, \quad \text{(B.74)}$$

$$P' \underset{(B.12)}{=} (1+\tau_C)(1-\rho\Delta)\omega = (1+\tau_C)(1-\rho\Delta)D(\theta,\epsilon,\Delta)P \quad \text{(B.75)}$$

Combining the closed form solution for the value function (B.71) with the Bellman equation (B.70), and substituting in for c and P' with the expressions above, we obtain

$$\bar{V}(\theta, X^*, \Delta) + \frac{1}{1-\alpha_\Delta}\log(P)\Delta = \int_{\epsilon>0}\Bigg(\log((\rho+\gamma+\rho\gamma\Delta)D(\theta,\epsilon,\Delta)P)\Delta + \alpha_\Delta$$
$$\times\Bigg\{[1-\lambda_\theta\Delta(1-g(\theta))]\Big(\bar{V}(\theta, X^*, \Delta) + \frac{1}{1-\alpha_\Delta}\log((1+\tau_C)(1-\rho\Delta)D(\theta,\epsilon,\Delta)P)\Delta\Big)$$
$$+\lambda_\theta\Delta(1-g(\theta))\Big(\bar{V}(1-\theta, X^*, \Delta) + \frac{1}{1-\alpha_\Delta}\log((1+\tau_C)(1-\rho\Delta)D(1-\theta,\epsilon,\Delta)P)\Delta\Big)\Bigg\}\Bigg)\mathrm{d}H(\epsilon). \quad \text{(B.76)}$$

This simplifies to the following expression for $\bar{V}(\theta, X^*, \Delta)$:

$$\bar{V}(\theta, X^*, \Delta) = \int_{\epsilon>0}\Bigg(\frac{\Delta}{1-\alpha_\Delta[1-\Delta\lambda_\theta(1-g(\theta))]}\Bigg\{\log((\rho+\gamma+\rho\gamma\Delta)D(\theta,\epsilon,\Delta))$$
$$+\alpha_\Delta[1-\lambda_\theta\Delta(1-g(\theta))]\frac{1}{1-\alpha_\Delta}\log((1+\tau_C)(1-\rho\Delta)D(\theta,\epsilon,\Delta))$$
$$+\alpha_\Delta\lambda_\theta\Delta(1-g(\theta))\Big(\bar{V}(1-\theta, X^*, \Delta)\frac{1}{\Delta} + \frac{1}{1-\alpha_\Delta}\log((1+\tau_C)(1-\rho\Delta)D(1-\theta,\epsilon,\Delta))\Big)\Bigg\}\Bigg)\mathrm{d}H(\epsilon). \quad \text{(B.77)}$$

The above equation (B.77) allows us the write the following system of equations:

$$\bar{V}(\theta, X^*, \Delta) = E(\theta,\Delta) + F(\theta,\Delta)\bar{V}(1-\theta, X^*, \Delta) \quad \text{(B.78)}$$
$$\bar{V}(1-\theta, X^*, \Delta) = E(1-\theta,\Delta) + F(1-\theta,\Delta)\bar{V}(\theta, X^*, \Delta), \quad \text{(B.79)}$$

where

$$E(\theta,\Delta) := \int_{\epsilon>0} \frac{\Delta}{1-\alpha_\Delta[1-\lambda_\theta\Delta(1-g(\theta))]}\Bigg\{ \log((\rho+\gamma+\rho\gamma\Delta)D(\theta,\epsilon,\Delta))$$
$$+ \alpha_\Delta[1-\lambda_\theta\Delta(1-g(\theta))]\frac{1}{1-\alpha_\Delta} \log((1+\tau_C)(1-\rho\Delta)D(\theta,\epsilon,\Delta))$$
$$+ \alpha_\Delta\lambda_\theta\Delta(1-g(\theta))\frac{1}{1-\alpha_\Delta} \log((1+\tau_C)(1-\rho\Delta)D(1-\theta,\epsilon,\Delta)\Bigg\}\mathrm{d}H(\epsilon),$$

$$F(\theta,\Delta) := \frac{\alpha_\Delta\lambda_\theta\Delta(1-g(\theta))}{1-\alpha_\Delta[1-\lambda_\theta\Delta(1-g(\theta))]}. \tag{B.80}$$

Hence, we obtain

$$\bar{V}(\theta,X^*,\Delta) = \frac{1}{1-F(\theta,\Delta)F(1-\theta,\Delta)}\Big(E(\theta,\Delta)+F(\theta,\Delta)E(1-\theta,\Delta)\Big) \tag{B.81}$$

Using l'Hopital's rule:

$$\lim_{\Delta\to 0} F(\theta,\Delta) = \frac{\lambda_\theta(1-g(\theta))}{\rho+\gamma+\lambda_\theta(1-g(\theta))}. \tag{B.82}$$

Hence, we obtain the following limit expression:

$$\frac{1}{1-\lim_{\Delta\to 0}F(\theta,\Delta)\lim_{\Delta\to 0}F(1-\theta,\Delta)} = \frac{1}{1-\frac{\lambda_\theta(1-g(\theta))}{\rho+\gamma+\lambda_\theta(1-g(\theta))}\frac{\lambda_\theta(1-g(1-\theta))}{\rho+\gamma+\lambda_\theta(1-g(1-\theta))}}$$
$$= \frac{(\rho+\gamma+\lambda_\theta(1-g(\theta)))\,(\rho+\gamma+\lambda_\theta(1-g(1-\theta)))}{(\rho+\gamma)\,(\rho+\gamma+\lambda_\theta)}$$

Now, we consider the following limit:

$$\lim_{\Delta\to 0}\int_{\epsilon>0}\frac{1}{\Delta}\log\left((1+\tau_C)D(\theta,\epsilon,\Delta)\right)dH(\epsilon)$$
$$= \lim_{\Delta\to 0}\int_{\epsilon>0}\frac{1}{\Delta}\log\Big([\phi(\epsilon-1)(1+\tau_C)+(r_E\theta-r_F)\Delta(1-\tau_K)]\,\hat{k}_E(\theta)+(1+\tilde{R}_F\Delta)\Big)\,dH(\epsilon)$$
$$= \lim_{\Delta\to 0}\frac{1}{\Delta}\int_\xi \log\Bigg(\Bigg[\phi(1-\underline{\epsilon})(1+\tau_C)\left(\exp\left(\varphi\sqrt{\Delta}\xi-\frac{\varphi^2\Delta}{2}\right)-1\right)+(r_E\theta-r_F)\Delta(1-\tau_K)\Bigg]\hat{k}_E(\theta)$$
$$+(1+\tilde{R}_F)\Bigg)d\Phi(\xi)$$

We now take the Taylor expansion of the integrand in units of $\sqrt{\Delta}$ around the point

$\sqrt{\Delta} = 0$. To do this let:

$$
\begin{aligned}
a &= \phi(1-\underline{\epsilon})(1+\tau_C)\hat{k} \\
b &= \varphi\xi \\
c &= \frac{\varphi^2}{2} \\
d &= (r_E\theta - r_F)(1-\tau_K)\hat{k} + \tilde{R}_F
\end{aligned}
$$

Then, we can write the integrand as $f(\sqrt{\Delta})$, where

$$f(\sqrt{\Delta}) = \log(a(\exp(b\Delta - c\Delta) - 1) + d\Delta + 1)$$

and so

$$
\begin{aligned}
f'(\sqrt{\Delta}) &= \frac{a(b-2c\sqrt{\Delta})\exp(b\sqrt{\Delta}-c\Delta) + 2d\sqrt{\Delta}}{a(\exp(b\sqrt{\Delta}-c\Delta)-1) + d\Delta + 1} \\
f''(\sqrt{\Delta}) &= (a(\exp(b\sqrt{\Delta}-c\Delta)-1) + d\Delta + 1)^{-2} \times \\
&\quad \Big\{ (a(\exp(b\sqrt{\Delta}-c\Delta)-1) + d\Delta + 1)\exp(b\sqrt{\Delta}-c\Delta)(-2ac + a(b-2c\sqrt{\Delta})^2 + 2d) \\
&\quad - (a(b-2c\sqrt{\Delta})\exp(b\sqrt{\Delta}-c\Delta) + 2d\sqrt{\Delta})^2 \Big\}
\end{aligned}
$$

Then, the Taylor expansion of the integrand above around $\sqrt{\Delta} = 0$ is as follows:

$$
\begin{aligned}
&\varphi\xi\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}\sqrt{\Delta} + \frac{\Delta}{2}(\xi^2-1)\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}\varphi^2 + \Delta(r_E\theta - r_F)(1-\tau_K)\hat{k} \\
&+ \Delta\tilde{R}_F - \frac{\Delta}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}\varphi\xi\right)^2
\end{aligned}
$$

Integrating this over ξ and using that $E[\xi] = 0$ and $E[\xi^2] = 1$, we therefore conclude that:

$$
\begin{aligned}
&\lim_{\Delta\to 0}\int_{\epsilon>0}\frac{1}{\Delta}\log\left((1+\tau_C)D(\theta,\epsilon,\Delta)\right)dH(\epsilon) \\
&\qquad = \lim_{\Delta\to 0}\frac{1}{\Delta}\left(\Delta(r_E\theta - r_F)(1-\tau_K)\hat{k} + \Delta\tilde{R}_F - \frac{\Delta}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}\varphi\right)^2 + \mathcal{R}\right) \\
&\qquad = (r_E\theta - r_F)(1-\tau_K)\hat{k} + \tilde{R}_F - \frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}\varphi\right)^2
\end{aligned}
$$

where $\mathcal{R}$ are some terms of higher order in Δ.

Now, rearranging our expression for $E(\theta, \Delta)$ above:

$$E(\theta,\Delta) = \left(\frac{\Delta}{1-\alpha_\Delta[1-\lambda_\theta\Delta(1-g(\theta))]}\right)$$
$$\times\left\{\log\left(\frac{\rho+\gamma+\rho\gamma\Delta}{1+\tau_C}\right)+\Delta\left(\int_{\epsilon>0}\frac{1}{\Delta}\log((1+\tau_C)D(\theta,\epsilon,\Delta))\mathrm{d}H(\epsilon)\right)\right.$$
$$+\alpha_\Delta[1-\lambda_\theta\Delta(1-g(\theta))]\frac{\Delta}{1-\alpha_\Delta}\left(\frac{\log(1-\rho\Delta)}{\Delta}+\int_{\epsilon>0}\frac{1}{\Delta}\log((1+\tau_C)D(\theta,\epsilon,\Delta))\mathrm{d}H(\epsilon)\right)$$
$$\left.+\alpha_\Delta\lambda_\theta\Delta(1-g(\theta))\frac{\Delta}{1-\alpha_\Delta}\left(\frac{\log(1-\rho\Delta)}{\Delta}+\int_{\epsilon>0}\frac{1}{\Delta}\log((1+\tau_C)D(1-\theta,\epsilon,\Delta))\mathrm{d}H(\epsilon)\right)\right\},$$

Evaluating the limit as Δ goes to zero for each right hand side term in turn, we obtain:

$$\lim_{\Delta\to0}E(\theta,\Delta) = \left(\frac{1}{\gamma+\rho+\lambda_\theta(1-g(\theta))}\right)\left\{\log\left(\frac{\rho+\gamma}{1+\tau_C}\right)+0\right.$$
$$\left.+\frac{1}{\rho+\gamma}\left(-\rho+\lim_{\Delta\to0}\int_{\epsilon>0}\frac{1}{\Delta}\log((1+\tau_C)D(\theta,\epsilon,\Delta))\mathrm{d}H(\epsilon)\right)+0\right\}$$
$$=\left(\frac{1}{\gamma+\rho+\lambda_\theta(1-g(\theta))}\right)\left\{\log\left(\frac{\rho+\gamma}{1+\tau_C}\right)\right.$$
$$\left.-\frac{\rho}{\rho+\gamma}+\frac{1}{\rho+\gamma}\left((r_E\theta-r_F)(1-\tau_K)\hat{k}(\theta)+\tilde{R}_F-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(\theta)\varphi\right)^2\right)\right\}$$

Taking the limit as $\Delta\to0$ in equation (B.81) and substituting in for $\lim_{\Delta\to0}F(\theta,\Delta)$ and $\lim_{\Delta\to0}E(\theta,\Delta)$ derived above, we obtain:

$$\lim_{\Delta\to0}\bar{V}(\theta,X^*,\Delta)=\frac{1}{1-\lim_{\Delta\to0}F(\theta,\Delta)\lim_{\Delta\to0}F(1-\theta,\Delta)}\left(\lim_{\Delta\to0}E(\theta,\Delta)+\lim_{\Delta\to0}F(\theta,\Delta)\lim_{\Delta\to0}E(1-\theta,\Delta)\right)$$
$$=\frac{(\rho+\gamma+\lambda_\theta(1-g(\theta)))\,(\rho+\gamma+\lambda_\theta(1-g(1-\theta)))}{(\rho+\gamma)\,(\rho+\gamma+\lambda_\theta)}\left\{\left(\log\left(\frac{\rho+\gamma}{1+\tau_C}\right)+\frac{\tilde{R}_F-\rho}{\rho+\gamma}\right)\right.$$
$$\times\left(\frac{1}{\rho+\gamma+\lambda_\theta(1-g(\theta))}+\frac{\lambda_\theta(1-g(\theta))}{\rho+\gamma+\lambda_\theta(1-g(\theta))}\frac{1}{\rho+\gamma+\lambda_\theta(1-g(1-\theta))}\right)$$
$$+\frac{1}{\rho+\gamma+\lambda_\theta(1-g(\theta))}\frac{1}{\rho+\gamma}\left((r_E\theta-r_F)(1-\tau_K)\hat{k}(\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(\theta)\varphi\right)^2\right)$$
$$+\frac{\lambda_\theta(1-g(\theta))}{\rho+\gamma+\lambda_\theta(1-g(\theta))}\frac{1}{\rho+\gamma+\lambda_\theta(1-g(1-\theta))}\frac{1}{\rho+\gamma}$$
$$\left.\times\left((r_E(1-\theta)-r_F)(1-\tau_K)\hat{k}(1-\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(1-\theta)\varphi\right)^2\right)\right\}$$

Then,

$$(\gamma+\rho)\lim_{\Delta\to 0}\bar{V}(\theta,X^*,\Delta) = \log\left(\frac{\gamma+\rho}{1+\tau_C}\right) + \frac{\tilde{R}_F-\rho}{\gamma+\rho}$$
$$+\frac{1}{\gamma+\rho}\frac{\gamma+\rho+\lambda_\theta g(\theta)}{\gamma+\rho+\lambda_\theta}\left((r_E\theta-r_F)(1-\tau_K)\hat{k}(\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(\theta)\varphi\right)^2\right)$$
$$+\frac{1}{\gamma+\rho}\frac{\lambda_\theta g(1-\theta)}{\gamma+\rho+\lambda_\theta}\left((r_E(1-\theta)-r_F)(1-\tau_K)\hat{k}(1-\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(1-\theta)\varphi\right)^2\right)$$

Substituting the above expression for $\bar{V}(\theta,X^*,\Delta)$ into equation (B.72), we obtain the the continuous time entrepreneur value function as stated in equation B.69, concluding the derivation.

$$(\gamma+\rho)V(P,\theta,X^*) = \log\left(\frac{\gamma+\rho}{1+\tau_C}P\right) + \frac{\tilde{R}_F-\rho}{\gamma+\rho}$$
$$+\frac{1}{\gamma+\rho}\frac{\gamma+\rho+\lambda_\theta g(\theta)}{\gamma+\rho+\lambda_\theta}\left((r_E\theta-r_F)(1-\tau_K)\hat{k}(\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(\theta)\varphi\right)^2\right)$$
$$+\frac{1}{\gamma+\rho}\frac{\lambda_\theta g(1-\theta)}{\gamma+\rho+\lambda_\theta}\left((r_E(1-\theta)-r_F)(1-\tau_K)\hat{k}(1-\theta)-\frac{1}{2}\left(\phi(1-\underline{\epsilon})(1+\tau_C)\hat{k}(1-\theta)\varphi\right)^2\right)$$

B.5 Characterizing Partial Equilibrium Elasticities of Tax Changes

To obtain a characterization, it is necessary to precisely define and compute the derivative $\frac{\partial H}{\partial \tau_j}\big|_{\mathcal{R}=0,N}$ for each $H \in \{K; K_E; N\}$ and $j \in \{K; W; C\}$, since this appears in Definition 3. To fix ideas, we first discuss how to compute $\frac{\partial Y}{\partial \tau_K}\big|_{\mathcal{R}=0,N}$, that is the marginal effect of τ_K on the steady state value of Y, holding fixed pre-tax prices and N. Other partial elasticities can then be defined in a similar fashion.

We may formally define $\frac{\partial Y}{\partial \tau_K}\big|_{\mathcal{R}=0,N}$ using Proposition 2 and equations (18) and (20). First, we use equations (18) and (20), which defined post-tax prices in terms of pre-tax prices and tax rates. Since we are seeking partial equilibrium effects, we differentiate these equations with repect to τ_K, holding constant pre-tax prices to obtain:

$$\frac{\partial \tilde{r}_X}{\partial \tau_K}\Big|_{\mathcal{R}=0,N} = \frac{r_F-r_E}{1+\mathcal{R}} = (r_E - r_F)$$
$$\frac{\partial \tilde{R}_F}{\partial \tau_K}\Big|_{\mathcal{R}=0,N} = r_F - \delta$$

Then, we make use of Proposition 2 which characterized a steady state using 8 equations, with 11 variables (as well as some strict inequality constraints). Invoking the implicit function theorem, it follows that in the neighborhood of some initial steady state $\mathscr{S}$, we can write the steady state values of eight of the variables as continuously differentiable functions of the three variables $\tilde{r}_X$, $\tilde{R}_F$ and τ_C.[19] Therefore, we may treat K, Y, K_E and C as functions of $\tilde{r}_X$, $\tilde{R}_F$ and τ_C, and use the equations of Proposition 2 to compute the partial derivatives $\frac{\partial Y}{\partial \tilde{r}_X}$ and $\frac{\partial Y}{\partial \tilde{R}_F}$. We compute the partial equilibrium partial derivatives, such as $\frac{\partial Y}{\partial \tilde{r}_X}\big|_{\mathcal{R}=0,N}$, by using Proposition 2 in this way, except that we ignore that $\mathcal{R} = \mathcal{R}(K_E, K, N, \pi_0)$ and instead fix $\mathcal{R} = 0$, and we ignore the condition for agents' optimal occupational choice (24) and instead hold N fixed at the initial steady state value. Combining these partial derivatives with the values of $\frac{\partial \tilde{r}_X}{\partial \tau_K}$ and $\frac{\partial \tilde{R}_F}{\partial \tau_K}$ found above, we may write:

$$\frac{\partial Y}{\partial \tau_K}\big|_{\mathcal{R}=0,N} = -(r_E - r_F)\frac{\partial Y}{\partial \tilde{r}_X}\big|_{\mathcal{R}=0,N} - (r_F - \delta)\frac{\partial Y}{\partial \tilde{R}_F}\big|_{\mathcal{R}=0,N},$$

and so

$$e^Y_{\tau_K} := -\frac{1-\tau_K}{Y}\left((r_E - r_F)\frac{\partial Y}{\partial \tilde{r}_X}\big|_{\mathcal{R}=0,N} + (r_F - \delta)\frac{\partial Y}{\partial \tilde{R}_F}\big|_{\mathcal{R}=0,N}\right).$$

The same logic can be used to define $\frac{\partial H}{\tau_j}$ for any aggregate steady state variable H, and for $j \in \{K; W; C\}$. In particular, equations (18) and (20) imply that for any such H:

$$\frac{\partial H}{\partial \tau_K}\big|_{\mathcal{R}=0,N} = -(r_E - r_F)\frac{\partial H}{\partial \tilde{r}_X}\big|_{\mathcal{R}=0,N} - (r_F - \delta)\frac{\partial H}{\partial \tilde{R}_F}\big|_{\mathcal{R}=0,N} \tag{B.83}$$

$$\frac{\partial H}{\partial \tau_W}\big|_{\mathcal{R}=0,N} = -\frac{\partial H}{\partial \tilde{R}_F}\big|_{\mathcal{R}=0,N}, \tag{B.84}$$

and the derivatives $\frac{\partial H}{\partial \tilde{r}_X}\big|_{\mathcal{R}=0,N}$, $\frac{\partial H}{\partial \tilde{R}_F}\big|_{\mathcal{R}=0,N}$ and $\frac{\partial H}{\partial \tau_C}\big|_{\mathcal{R}=0,N}$ can all be defined using Proposition 2 and the implicit function theorem, as described above.

Using this approach, we now provide a characterization of the elasticities of Y, K_E and K with respect to taxes.

B.6 Effects of Tax Changes on Welfare

Since workers choose $\frac{a^N_{s+1}}{1+\tau_C}$ optimally each period, envelope theorem arguments imply that we may calculate the resulting change in welfare as if workers continue to choose the same level of $\frac{a^N_{s+1}}{1+\tau_C}$ each period irrespective of the tax change. Then, the worker's budget constraint implies that the tax change has an effect on welfare equivalent to increasing worker consumption by $\mathrm{d}c^N_s$ in each period s, where $\mathrm{d}c^N_s$ satisfies:

$$\mathrm{d}c^N_s = \mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \frac{a^N_s}{1+\tau_C}$$

where $\mathrm{d}\tilde{w}$ and $\mathrm{d}\tilde{R}_F$ are the change in $\tilde{w}$ and $\tilde{R}_F$ as a result of the tax change.

[19] This holds, provided the relevant Jacobian is invertible, which will be the case outside of knife-edge situations.

In such a case, the tax change increases the present value of the worker's lifetime resources by:

$$\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s \mathrm{d}c_s^N = \left(\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s\right)\left(\mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \frac{\mathcal{A}^N}{1+\tau_C}\right)$$

where

$$\mathcal{A}^N = \frac{\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s a_s^N}{\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s},$$

is the average value of the worker's discounted lifetime assets. Applying envelope theorem arguments further, the change in welfare from a small tax change must then be equivalent to the change in worker utility if the worker consumed all the extra resources $\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s \mathrm{d}c_s^N$ in the first period of their life, since on the margin, workers are indifferent about which period they consume each extra unit of lifetime resources they receive. That is, the change in welfare satisfies:

$$\mathrm{d}\mathcal{W} = \frac{1}{c_0^N}\left(\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s\right)\left(\mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \frac{\mathcal{A}^N}{1+\tau_C}\right),$$

where $\frac{1}{c_0^N}$ is the worker's marginal utility of consumption in the first year of her life.

To calculate the consumption equivalent welfare change Δ^N, note that multiplying worker consumption by $1+\Delta^N$ each period increases the present value of a newborn worker's lifetime consumption stream by:

$$\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s c_s^N \Delta^N = \sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s \tilde{w}\Delta^N,$$

where we used that the budget constraint implies that the lifetime present value of a worker's consumption must equal the lifetime present value of his earnings.

Applying envelope theorem arguments once more, a multiplication of worker consumption by $1+\Delta^N$ each period must have the same effect on worker welfare as increasing a worker's consumption by $\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s \tilde{w}\Delta^N$ in the first period of life and leaving his consumption constant thereafter. Then, the change Δ^N that results in a welfare change $\mathrm{d}\mathcal{W}$ must satisfy:

$$\left(\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s\right)\tilde{w}\Delta^N \frac{1}{c_0^N} = \mathrm{d}\mathcal{W}$$

Combining this with the equation for $\mathrm{d}\mathcal{W}$ above, we obtain a simple formula for Δ^N:

$$\tilde{w}\Delta^N = \mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \frac{\mathcal{A}^N}{1+\tau_C}.$$

B.7 Proof of Lemma ??

Formally, we define the government's optimization problem as seeking taxes $\tau_K^\star, \tau_C^\star, \tau_N^\star, \tau_W^\star$ and an allocation to achieve the supremum of worker steady state utility subject to the constraint that all aggregate variables must be consistent with the equations and inequalities in Proposition 2 – i.e. the allocation must be a steady state.

Since we are seeking a supremum to the government's problem, all the strict inequalities in Proposition 2 can be replaced with weak inequalities, since all the inequalities are continuous functions of the aggregate variables. After this replacement, all the constraints of the government's problem are either equalities or weak inequalities.

Consider a compact neighborhood of combinations of feasible taxes and allocations around the optimal taxes $\tau_K^\star, \tau_C^\star, \tau_N^\star, \tau_W^\star$. Such a compact set exists, since all the constraints of the government's problem are equalities and weak inequalities. Since the government's objective function is continuous over this compact set, it follows that, within this compact set, the maxium of the government's objective function must be attained at the tax rates $\tau_K^\star, \tau_C^\star, \tau_N^\star, \tau_W^\star$, by the Weierstrass theorem.

At the optimal tax rates $\tau_K^\star, \tau_C^\star, \tau_N^\star, \tau_W^\star$ and optimal allocation, it must be the case that $r_E^* > r_F^*$, since, by the Inada conditions on the production function, $r_E^* = \infty$, if $K_E = 0$, and if $K_E > 0$ then equation (**??**) implies that it must be the case that $r_E^* > r_F^*$, since $\tau_K^\star < 1$. Since $r_E^* > r_F^*$, it must also be the case that $K_F^* > 0$, since the Inada conditions imply $r_F^* = \infty$ otherwise.

Then, the only inequality constraints on the government's problem that may bind at the optimal allocation are that $\tilde{R}_F = (1 + \tau_C^*)(\tilde{r}_F^* - \tilde{p}^*) \geq \underline{r}$ and $\frac{1}{1+\tau_C^*} \geq \phi$. Since all the relevant functions are continuously differentiable, we may apply the Kuhn-Tucker theorem to deduce that the optimum must satisfy the Kuhn-Tucker first order conditions, with Lagrange multipliers on these latter two inequality constraints that may bind. Combining the steady state conditions in Proposition 2 to get an expression for worker's post tax utility as a function of K_E and K_F and differentiating yields the first order conditions in the Lemma.

□

C Data

To calibrate the entrepreneur's stake in the business, we use data from two sources: the Survey of Consumer Finances (SCF) and the (National) Survey of Small Business Finances (SSBF). Both surveys contain information regarding business ownership, with the difference that the first is a household survey, while the second is a survey of small businesses. We can identify in each of them groups of respondents that are in line with our notion of entrepreneurship. We use both sources as validation for our results.

The Survey of Consumer Finances is a a triennial cross-sectional survey of U.S. families which provides information on individual household portfolio composition, including investment in private firms. While the SCF was initially administered in 1983, it was not until 1989 that questions about business ownership were introduced. Therefore, we use all survey waves from 1989 until 2013. We restrict the sample to households who report owning a business in which they have an active management interest, and are between 25 and 65 years old. This represents, on average, 14.3% of the sample. If a household is an active participant in multiple businesses, we examine the average share across businesses.[20]

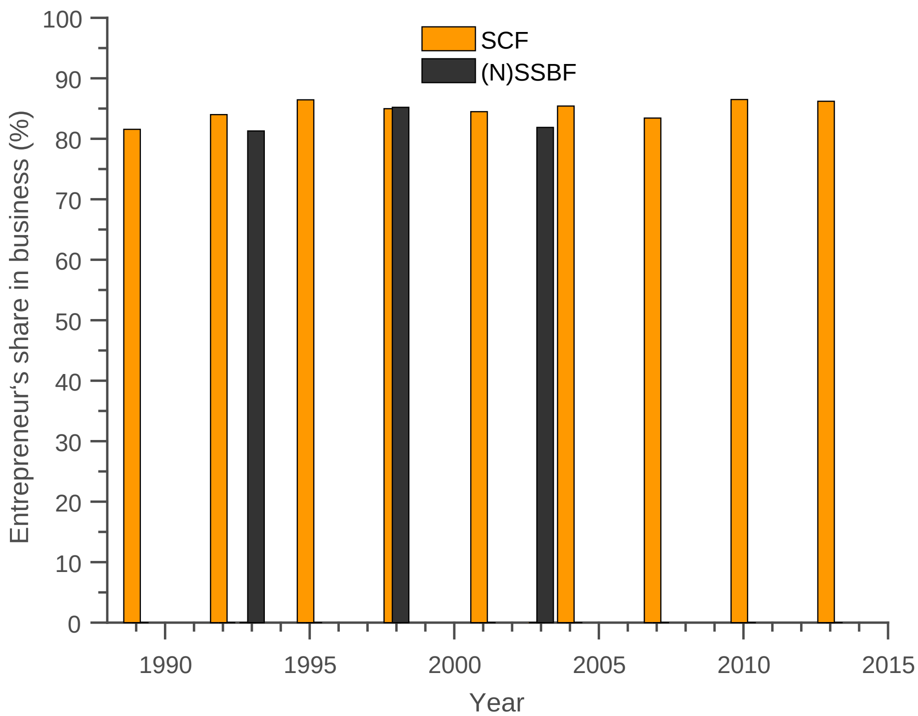
The (National) Survey of Small Business Finances collects information on private, non-financial, non-agricultural businesses in the U.S., with fewer than 500 employees. There are four surveys to date, but only the last three (1993, 1998 and 2003) collect ownership share information and are useful for our purposes. The surveys detail the demographic and financial characteristics of the firms and their principal shareholder.[21] Approximately 90% of these firms are managed by the principal shareholder. We apply the same sample restrictions as in the SCF.

Figure 2 displays the evolution of the ownership share over time. Both surveys indicate that ownership is highly concentrated, entrepreneurs holding, on average, 84% of their firm's equity. In particular, the average share is 85% in SCF and 83% in (N)SSBF. Ownership rates are very stable not only across surveys, but also across the time horizon we consider. For this reason, for our calibration exercise we work with their average over time and surveys, which is 84%.

[20]We obtain similar results if we only focus on the business in which the household has the largest investment.

[21]In 2003 information was collected for up to three owners. We only focus on the main owner, i.e. the one with the largest ownership in the business.

Figure 2: Ownership Share in the U.S.



Notes: The orange bars show the average share that entrepreneurs in SCF own in their business. The black bars show the average share of small businesses in the (N)SSBF that is owned by the principal shareholder.