# Optimal Development Policies with Financial Frictions\*

Oleg Itskhoki itskhoki@princeton.edu Benjamin Moll moll@princeton.edu

July 24, 2018

#### Abstract

Is there a role for governments in emerging countries to accelerate economic development by intervening in product and factor markets? To address this question, we study optimal dynamic Ramsey policies in a standard growth model with financial frictions. The optimal policy intervention involves pro-business policies like suppressed wages in early stages of the transition, resulting in higher entrepreneurial profits and faster wealth accumulation. This in turn relaxes borrowing constraints in the future, leading to higher labor productivity and wages. In the long run, optimal policy reverses sign and becomes pro-worker. In a multi-sector extension, optimal policy subsidizes sectors with a latent comparative advantage and under certain circumstances involves a depreciated real exchange rate. Our results provide an efficiency rationale, but also identify caveats, for many of the development policies actively pursued by dynamic emerging economies.

<sup>\*</sup>We are particularly grateful to Mike Golosov for many stimulating discussions. We also thank Iván Werning, Joe Kaboski and Sangeeta Pratap for insightful discussions, and Mark Aguiar, Marco Bassetto, Paco Buera, Steve Davis, Emmanuel Farhi, Gene Grossman, Chang-Tai Hsieh, Keyu Jin, Nobu Kiyotaki, Guido Lorenzoni, Juan-Pablo Nicolini, Gerard Padró i Miquel, Rob Shimer, Yongs Shin, Gianluca Violante, Yong Wang and seminar participants at various institutions and conferences for very helpful comments. Kevin Lim, Mathis Maehlum, Dmitry Mukhin and Max Vogler provided outstanding research assistance. Moll thanks the Kauffman Foundation for financial support.

## 1 Introduction

Is there a role for governments in emerging countries to accelerate economic development by intervening in product and factor markets? If so, which policies should they adopt? To answer these questions, we study optimal policy interventions in a standard growth model with financial frictions. In our framework, forward-looking heterogeneous producers face borrowing (collateral) constraints that result in capital misallocation and depressed productivity. This framework is, therefore, similar to the one commonly adopted in the macro-development literature to study the relationship between financial development and aggregate productivity (see e.g. Banerjee and Duflo, 2005; Song, Storesletten, and Zilibotti, 2011; Buera and Shin, 2013). Our paper is the first to study the optimal Ramsey policies in such an environment along with their implications for a country's development dynamics.

Our main result is that the optimal policy involves interventions in both product and factor markets, yet the direction of these interventions is different for developing and developed countries, defined in terms of the level of their financial wealth relative to the steady state. In particular, in the initial phase of transition, when entrepreneurs are undercapitalized, optimal policies are *pro-business* in the sense of shifting resources towards entrepreneurs. Once the economy comes close enough to the steady state, where entrepreneurs are well capitalized, optimal policy switches to being *pro-worker*. Hence, optimal policy is *stage-dependent*.<sup>1</sup> In the case of the labor market, it is optimal to increase labor supply and suppress equilibrium wages in early stages of development, and restrict labor supply later on.<sup>2</sup> Greater labor supply and suppressed wages increase entrepreneurial profits and accelerate wealth accumulation. This, in turn, makes future financial constraints less binding, resulting in greater labor productivity and higher wages.

From a more positive perspective, we are motivated by the observation that many successful emerging economies pursue active development and industrial policies and in particular policies that appear to favor businesses. A widespread example of such policies is the suppression or subsidization of factor prices. For example, South Korea in the 1970s imposed an official upper limit on the growth of real wages,<sup>3</sup> and we discuss other exam-

<sup>&</sup>lt;sup>1</sup>In "Optimal Development Policies: Lessons from Experience," Tinbergen (1984, p. 112) writes: "Development policy' was the name given to the new endeavours whose ideal was to raise the standard of living in the way best possible in the prevailing circumstances. [...] What is still needed is an optimal development policy within ever-changing surroundings." Our analysis of dynamic optimal development policies is an attempt to provide one possible answer to this call.

<sup>&</sup>lt;sup>2</sup>This reduced labor supply and the resulting increase in wages are reminiscent of a *labor union* allocation.

<sup>&</sup>lt;sup>3</sup>South Korea's Economic Planning Board directed firms to keep nominal wage growth below 80 percent of the sum of inflation and aggregate productivity growth, which resulted in real wage growth lagging behind

ples at the end of this introduction. From a neoclassical perspective such policies are, of course, unambiguously detrimental. In this paper we argue that some of these policies may instead be beneficial, particularly at early stages of development. Our result on the stagedependence of optimal policies provides an efficiency rationale for the different labor market institutions adopted by emerging Asian and developed European countries, without relying on differences in preferences or political systems.

We tackle the design of optimal policy using a tractable workhorse macro-development model, which allows us to obtain sharp analytical characterizations. Our baseline economy is populated by two types of agents: a continuum of heterogeneous entrepreneurs and a continuum of representative workers. Entrepreneurs differ in their wealth and their productivity, and borrowing constraints limit the extent to which capital is reallocated from wealthy to productive individuals. In the presence of financial frictions, productive entrepreneurs make positive profits, and then optimally choose how much of these to consume and how much to retain for wealth accumulation. Workers decide how much labor to supply to the market and how much to save. Our baseline framework builds on Moll (2014) and makes use of the insight that heterogeneous agent economies remain tractable if individual production functions feature constant returns to scale.<sup>4</sup> Section 2 lays out the structure of the economy and characterizes the decentralized laissez-faire equilibrium. As a result of financial frictions, marginal products of capital are not equalized across agents, and constrained entrepreneurs obtain a higher rate of return than that available to workers. It is this differential in rates of return that is exploited by the policy interventions we consider.

In Section 3, we introduce various tax instruments into this economy and study the optimal Ramsey policies given the available set of instruments. We consider the problem of a benevolent planner subject to the same financial frictions present in the decentralized economy. We first consider the case with a subset of tax instruments, which effectively allow the planner to manipulate worker savings and labor supply decisions, and then show how the results generalize to cases with a much larger set of instruments, which in particular includes credit subsidies to firms (entrepreneurs). As already mentioned, the optimal policy intervention increases labor supply in the initial phase of transition, when entrepreneurs are undercapitalized, and reduces labor supply once the economy comes close enough to the steady state. We show in Section 3.3 that it remains optimal to distort labor supply in this

productivity growth (see Kim and Topel, 1995; Kim and Leipziger, 1997). Labor unions were also restricted. On the anecdotal side, president Park Chung Hee in his annual national address declared 1965 a "year to work," and twelve months later, he humorlessly named 1966 a "year of harder work" (Schuman, 2010).

<sup>&</sup>lt;sup>4</sup>By adopting specific distributional assumptions, we gain additional tractability essential for our dynamic optimal policy analysis and the various extensions we consider later in the paper.

fashion even when credit or capital subsidies are available, which are arguably more direct instruments for targeting the underlying inefficiency. Furthermore, in Section 3.4, we show that this policy remains optimal even when workers are finitely lived as in Blanchard (1985) and Yaari (1965), face borrowing constraints, and when the planner is present-biased in favor of current generations.

While our benchmark analysis focuses on a labor supply subsidy for concreteness, there are of course many equivalent ways of implementing the optimal allocation, including non-tax market regulation which is widespread in practice, as we discuss in Section 3. The common feature of optimal policies is that, in the short run, they make workers work more even though wages paid by firms are low. Perhaps surprisingly, we show that such *pro-business* development policies are optimal even when the planner puts zero weight on the welfare of entrepreneurs. Indeed, the planner finds it optimal to hurt workers in the short run so as to reward them with higher wages and shorter work hours in the medium and long run. An alternative way of thinking about this result is that the labor supply decision of workers involves a dynamic *pecuniary externality* (see Greenwald and Stiglitz, 1986): workers do not internalize the fact that working more leads to faster wealth accumulation by entrepreneurs and higher wages in the future. The planner corrects this using a Pigouvian subsidy.<sup>5</sup>

In order to obtain analytical results, we adopt a number of assumptions which allow for tractable aggregation of the economy and result in a simple characterization of equilibrium dynamics under various government policies. The most important among these are constant returns in production, *iid* productivity shocks drawn from an unbounded Pareto distribution, and no financial constraints on workers. In Section 4, to relax these assumptions, we give up analytical tractability and extend the model to a richer quantitative environment featuring a time-varying joint distribution of wealth and productivity as a state variable, making optimal policy analysis a challenging task at the computational frontier. This allows us to examine the robustness of our results as well as to gauge the quantitative importance of both the optimal and various suboptimal policies. We find that the optimal policies are stage-dependent as in our analytical results and can lead to considerable welfare improvements. Our quantitative results therefore confirm our main message that pro-business policies are stage especially important for growth at earlier stages of development, and that such policies can be welfare-improving even from workers' perspective.

<sup>&</sup>lt;sup>5</sup>We show that a reduced form of our model is mathematically equivalent to a setup in which production is subject to a learning-by-doing externality, whereby working more today increases future productivity, as in Krugman (1987) and Young (1991). While reduced forms are similar, the economies are structurally different: the dynamic externality in our framework is a pecuniary one, stemming from the presence of financial frictions and operating via misallocation of resources, rather than a technological one.

In Section 5, we take advantage of the tractability of our framework and extend the model to multiple tradable and non-tradable sectors. This allows us to study the optimal industrial policies and address a number of popular policy issues, such as promotion of comparative advantage sectors (see e.g. Lin, 2012) and optimal exchange rate policy (see e.g. Rodrik, 2008). We show, for example, that financial frictions create a wedge between the short-run and long-run (*latent*) comparative advantage of a country, and that the optimal policy tilts the allocation of resources towards the latent comparative advantage sectors, thereby speeding up the transition.<sup>6</sup> Next, we identify circumstances under which optimal policy involves a depreciated real exchange rate. Lastly, we develop an extension with overlapping cohorts of entrepreneurs and show that optimal policy requires age-dependent subsidies akin to the popular policy of *infant industry protection*.

**Empirical relevance** There exists a large number of historical accounts that the sort of pro-business policies prescribed by our normative analysis have been used in countries with successful development experiences. In a companion Online Appendix B, we discuss in detail development policies in seven East Asian countries that have experienced episodes of rapid growth: Japan, Korea, Taiwan, Malaysia, Singapore, Thailand and China.<sup>7</sup> Typical policies include the suppression of wages and intermediate input prices. Another commonly observed policy is some form of subsidized credit to particular sectors or firms, often conditional on their export status. Many of these policies are reminiscent of the normative prescriptions in our theoretical analysis for economies in the early stages of development. In practice, such policies were frequently imposed for reasons other than development, e.g. due to political, ideological or rent-seeking considerations (see e.g. Harrison and Rodríguez-Clare, 2010). Yet, our analysis suggests that successful growth episodes may have occurred not despite but, at least in part, due to the adoption of such policies.

From a more historical perspective, Feinstein (1998) and Voth (2001) provide evidence that the rapid economic growth in 18th-century Britain was in part due to reduced labor and land prices as well as long work hours. Ventura and Voth (2015) argue that this was caused by expanding government borrowing which crowded out unproductive agricultural investment and reduced factor demand by this declining sector. Lower factor prices, in turn, increased profits in the new industrial sectors, allowing the capitalists in these sectors to build up wealth, which in the absence of an efficient financial system was the major source of investment. This historical account resonates well with the mechanics of our model.

<sup>&</sup>lt;sup>6</sup>Such policies have even been embraced by the World Bank: their Growth Commission (2008) report argues that export promotion policies may be beneficial, at least as long as they are only temporary.

<sup>&</sup>lt;sup>7</sup>Appendix B is available at http://www.princeton.edu/~itskhoki/papers/FinFrictionsDevoPolicy\_AppendixB.pdf.

**Related literature** Our paper is related to the large theoretical literature studying the role of financial market imperfections in economic development, and in particular the more recent literature relating financial frictions to aggregate productivity. We contribute to this literature by studying optimal Ramsey policies and the resulting implications for a country's transition dynamics in both a one-sector and a multi-sector environment.<sup>8</sup>

In related work, Caballero and Lorenzoni (2014) analyze the Ramsey-optimal response to a cyclical demand shock in a two-sector small open economy with financial frictions in the tradable sector. In both papers financial frictions give rise to a pecuniary externality, which justifies a policy intervention that distorts the allocation of resources across sectors. But the focus of the two papers is different: ours studies long-run development policies whereas theirs studies cyclical policies.<sup>9</sup> Another closely related paper is Song, Storesletten, and Zilibotti (2014) who study the effects of capital controls and policies regulating interest rates and the exchange rate. Their positive analysis shows that, in China, such policy interventions may have compressed wages and increased the wealth of entrepreneurs, relaxing the borrowing constraints of private firms, just like in our framework. Relative to their paper, our normative analysis shows that policies leading to compressed wages not only foster productivity growth but may, in fact, be *optimal* in the sense of maximizing welfare.

The general idea that different policies may be appropriate at different stages of a country's development has previously been explored by Acemoglu, Aghion, and Zilibotti (2006). They argue that countries far behind the frontier should adopt investment subsidies and other policies that increase firm profits and then, as they get closer to the frontier, switch to policies supporting innovation and selection. In their framework, the rationale for such policies is a Schumpeterian appropriability effect. In our framework, in contrast, the laissez-faire

<sup>&</sup>lt;sup>8</sup>In addition to the papers cited in the beginning of the introduction, see also Banerjee and Newman (1993), Galor and Zeira (1993), Aghion and Bolton (1997), Jeong and Townsend (2007), Erosa and Hidalgo-Cabrillana (2008), Caselli and Gennaioli (2013), Amaral and Quintin (2010), Buera, Kaboski, and Shin (2011), Midrigan and Xu (2014) and the recent surveys by Matsuyama (2008) and Townsend (2010). These papers are part of a growing literature exploring the macroeconomic effects of micro distortions (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). The modeling of financial frictions in the paper also follows the tradition in the recently burgeoning macro-finance literature (Kiyotaki and Moore, 1997; Brunnermeier, Eisenbach, and Sannikov, 2013). A few papers in this literature evaluate the effects of various policies, including Banerjee and Newman (2003), Buera, Moll, and Shin (2013), Buera and Nicolini (2017) and Buera, Kaboski, and Shin (2012), but none study Ramsey-optimal policies. There is an even larger *empirical* literature showing the importance of finance for development (see Levine, 2005, for a survey).

<sup>&</sup>lt;sup>9</sup>See also Angeletos, Collard, Dellas, and Diba (2013) and Bacchetta, Benhima, and Kalantzis (2014) for related Ramsey problems and Michelacci and Quadrini (2009) for a related study of optimal long-term contracts between employers and employees. A related strand of work emphasizes a different type of pecuniary externality that operates through prices in borrowing constraints (see e.g. Lorenzoni, 2008; Jeanne and Korinek, 2010; Bianchi, 2011). Yet another type of pecuniary externality is analyzed in the earlier work on the "big push" (e.g., see Murphy, Shleifer, and Vishny, 1989).

equilibrium is suboptimal due to the presence of financial frictions.

In terms of methodology, we follow the dynamic public finance literature and study a Ramsey problem (see e.g. Barro, 1979; Lucas and Stokey, 1983). The environment we study is similar to Judd (1985) and Straub and Werning (2014) in that it features a distributional conflict between capitalists and workers, but with the difference that capitalists are heterogeneous and face financial frictions and incomplete markets.<sup>10</sup> Our work differs from the classical Ramsey taxation literature in that we study optimal policy intervention in the presence of financial frictions, rather than the optimal financing of an exogenously given stream of government expenditure or optimal debt management.

### 2 An Economy with Financial Frictions

In this section we describe our baseline one-sector small open economy. We extend our analysis to a closed economy in Appendix A4 and to a multi-sector economy with tradable and non-tradable sectors in Section 5.<sup>11</sup> Our goal is to develop a model of transition dynamics with financial frictions in which we can analyze optimal government interventions. This goal motivates adopting a number of assumptions which allow for tractable aggregation of the economy and result in a simple characterization of equilibrium dynamics under various government policies. We later relax many of these assumptions in Section 4.

The economy is set in continuous time with an infinite horizon and no aggregate shocks, so as to focus our analysis on the properties of the transition paths. There are two types of agents—workers and entrepreneurs—and we start by describing their problems in turn. We then characterize the aggregate relationships and properties of the decentralized equilibrium.

#### 2.1 Workers and entrepreneurs

A representative worker (household) in the economy has preferences given by

$$\int_0^\infty e^{-\rho t} u\big(c(t), \ell(t)\big) \mathrm{d}t,\tag{1}$$

where  $\rho$  is the discount rate, c is consumption, and  $\ell$  is market labor supply. We assume that  $u(\cdot)$  is increasing and concave in its first argument and decreasing and convex in its second

<sup>&</sup>lt;sup>10</sup>See Aiyagari (1995) and Shin (2006) for related analyses of Ramsey problems in environments with idiosyncratic risk and incomplete markets, but without collateral constraints.

<sup>&</sup>lt;sup>11</sup>Appendix A contains additional extensions, as well as detailed derivations and proofs for the baseline analysis, and can be found at http://www.princeton.edu/~itskhoki/papers/FinFrictionsDevoPolicy\_AppendixA.pdf.

argument, with a positive and finite Frisch elasticity of labor supply (see Appendix A1.1). Where it leads to no confusion, we drop the time index t.

Households take the market wage w(t) as given as well as the price of the consumption good, which we choose as the numeraire. They borrow and save using non-state-contingent bonds, which pay the risk-free interest rate  $r(t) \equiv r^*$ , and hence face the flow budget constraint

$$c + \dot{b} \le w\ell + r^*b,\tag{2}$$

where b(t) is the household asset position. The solution to the household problem satisfies a standard Euler equation and a static optimality (labor supply) condition. In Section 3.4, we extend our analysis to an environment with overlapping generations of finitely-lived households that also face borrowing constraints.

The economy is also populated by a unit mass of entrepreneurs who produce the homogeneous tradable good. Entrepreneurs are heterogeneous in their wealth a and productivity z, and we denote their joint distribution at time t by  $\mathcal{G}_t(a, z)$ . In each time period of length  $\Delta t$ , entrepreneurs draw a new productivity from a Pareto distribution  $G_z(z) = 1 - z^{-\eta}$  with shape parameter  $\eta > 1$ , where a smaller  $\eta$  corresponds to a greater heterogeneity of the productivity draws. We consider the limit economy in which  $\Delta t \to 0$ , so we have a continuous-time setting in which productivity shocks are *iid* over time.<sup>12</sup> Appendix A4.1 generalizes our *qualitative* results to an environment with a persistent productivity process, while Section 4 considers a *quantitative* version of the model with decreasing returns to scale and a diffusion process for idiosyncratic productivities, which render the model analytically intractable.

Entrepreneurs have preferences

$$\mathbb{E}_0 \int_0^\infty e^{-\delta t} \log c_e(t) \mathrm{d}t,\tag{3}$$

where  $\delta$  is their discount rate. Each entrepreneur owns a private technology which can combine k units of capital and n units of labor to produce

$$A(t)(zk)^{\alpha}n^{1-\alpha} \tag{4}$$

units of output, where  $\alpha \in (0, 1)$ , and A(t) is aggregate productivity, which follows an exogenous path. Entrepreneurs hire labor in a competitive labor market at wage w(t) and hire capital in a capital rental market at rental rate  $r(t) \equiv r^*$ .

<sup>&</sup>lt;sup>12</sup>Moll (2014) shows that an *iid* process in continuous time can be obtained by considering a limit of a mean-reverting process as the speed of mean reversion goes to infinity. In addition, we assume a law of large numbers so the share of entrepreneurs experiencing any particular sequence of shocks is deterministic.

Entrepreneurs face collateral constraints:

$$k \le \lambda a,\tag{5}$$

where  $\lambda \geq 1$  is an exogenous parameter, which captures the degree of financial development, from self-financing when  $\lambda = 1$  to perfect capital markets as  $\lambda \to \infty$ . By placing a restriction on an entrepreneur's leverage ratio k/a, the constraint captures the common prediction from models of limited commitment that the amount of capital available to an entrepreneur is limited by her personal wealth and that production requires a certain minimum *skin in the game*. Banerjee and Duflo (2005) survey evidence on the importance of such constraints for developing countries. The particular formulation of the constraint in (5) is analytically convenient and allows us to derive results in closed form.<sup>13</sup>

An entrepreneur's wealth evolves according to

$$\dot{a} = \pi(a, z) + r^* a - c_e, \tag{6}$$

where  $\pi(a, z)$  are her profits

$$\pi(a,z) \equiv \max_{\substack{n \ge 0, \\ 0 \le k \le \lambda a}} \left\{ A(zk)^{\alpha} n^{1-\alpha} - wn - r^*k \right\}.$$
(7)

Maximizing out the choice of labor n, profits are linear in capital k. It follows that the optimal capital choice is at a corner: it is zero for entrepreneurs with low productivity, and the maximal amount allowed by the collateral constraints,  $\lambda a$ , for those with high enough productivity. We assume that along all transition paths considered, there always exist entrepreneurs with productivity low enough that they choose to be inactive. In this case, the solution to (7) admits the following characterization (see Appendix A1.2):

<sup>&</sup>lt;sup>13</sup>Following the literature, we model financial frictions as the interaction between incomplete markets and collateral constraints, both exogenously imposed. The constraint can be derived from a limited commitment problem, in which an entrepreneur can steal a fraction  $1/\lambda$  of rented capital k, and lose her wealth a as a punishment (see Kiyotaki and Moore, 1997; Banerjee and Newman, 2003; Buera and Shin, 2013). As shown in Moll (2014), the constraint could be generalized in a number of ways at the expense of some extra notation. In particular, the maximum leverage ratio  $\lambda$  can depend on the interest rate, wages and other aggregate variables, or evolve over time. In addition, the financing friction may also extend to working capital needed to cover an entrepreneur's wage bill. What *is* crucial is that the collateral constraint (5) is linear in wealth and static (ruling out dynamic incentive contracts as e.g. in Kehoe and Levine, 2001). Di Tella and Sannikov (2016) provide a microfoundation to such constraints in a dynamic environment with hidden savings.

**Lemma 1** Factor demands and profits are linear in wealth and can be written as:

$$k(a,z) = \lambda a \cdot \mathbf{1}_{\{z \ge \underline{z}\}},\tag{8}$$

$$n(a,z) = \left[ (1-\alpha)A/w \right]^{1/\alpha} zk(a,z),$$
(9)

$$\pi(a,z) = \left[\frac{z}{\underline{z}} - 1\right] r^* k(a,z),\tag{10}$$

where the productivity cutoff  $\underline{z}$  satisfies:

$$\alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} A^{1/\alpha} \underline{z} = r^*.$$
(11)

Marginal entrepreneurs with productivity  $\underline{z}$  break even and make zero profits, while entrepreneurs with productivity  $z > \underline{z}$  receive Ricardian rents given by (10), which depend on both their productivity edge and the scale of operation determined by their wealth through the collateral constraint. Entrepreneurs' labor demand depends on both their productivity and their capital choice, with marginal products of labor equalized across active entrepreneurs. At the same time, the choice of capital among active entrepreneurs is shaped by the collateral constraint, which depends only on their assets and not on their productivity. Therefore, entrepreneurs with higher productivity z have a higher marginal product of capital, reflecting the misallocation of resources in the economy. The corner solution for the choice of capital in (8) is a consequence of the constant returns to scale assumption, which we relax in Section 4.

Finally, entrepreneurs choose consumption and savings to maximize (3) subject to (6) and (10). Under our assumption of log utility, combined with the linearity of profits in wealth, there exists an analytic solution to their consumption policy function,  $c_e = \delta a$ , and therefore the evolution of entrepreneurs' wealth satisfies (see Appendix A1.3):

$$\dot{a} = \pi(a, z) + (r^* - \delta)a.$$
 (12)

This completes our description of workers' and entrepreneurs' individual behavior.

### 2.2 Aggregation and equilibrium

We start by describing a number of useful equilibrium relationships. Aggregating (8) and (9) over all entrepreneurs, we obtain the aggregate capital and labor demand:

$$\kappa = \lambda x \underline{z}^{-\eta},\tag{13}$$

$$\ell = \frac{\eta}{\eta - 1} \left[ (1 - \alpha) A/w \right]^{1/\alpha} \lambda x \underline{z}^{1-\eta}, \tag{14}$$

where  $x(t) \equiv \int a d\mathcal{G}_t(a, z)$  is aggregate entrepreneurial wealth.<sup>14</sup> Note that we have made use of the assumption that productivity shocks are *iid* over time which implies that, at each point in time, wealth *a* and productivity *z* are independent in the cross-section of entrepreneurs. Intuitively, the aggregate demand for capital in (13) equals the aggregate leveraged wealth of the entrepreneurs  $\lambda x$  times the fraction of active entrepreneurs  $\underline{z}^{-\eta} = \mathbb{P}\{z \geq \underline{z}\}$ , as follows from the Pareto productivity distribution.

Aggregate output in the economy can be characterized by a production function:

$$y = Z \kappa^{\alpha} \ell^{1-\alpha}$$
 with  $Z \equiv A \left(\frac{\eta}{\eta-1} \underline{z}\right)^{\alpha}$ , (15)

where Z is the endogenous aggregate total factor productivity (TFP), which is a product of aggregate technology A and the average productivity of active entrepreneurs,  $\mathbb{E}\{z|z \ge z\} = \frac{\eta}{\eta-1}\overline{z}$ . Imposing labor market clearing, and using the aggregation results in (13)–(15) together with the productivity cutoff condition (11), we can characterize the equilibrium relationships in the frictional economy (see Appendix A1.2):

Lemma 2 (a) Equilibrium aggregate output satisfies:

$$y = y(x,\ell) \equiv \Theta \, x^{\gamma} \ell^{1-\gamma},\tag{16}$$

where  $\Theta \equiv \frac{r^*}{\alpha} \left[ \frac{\eta \lambda}{\eta - 1} \left( \frac{\alpha A}{r^*} \right)^{\eta/\alpha} \right]^{\gamma}$  and  $\gamma \equiv \frac{\alpha/\eta}{\alpha/\eta + (1 - \alpha)}.$ 

(b) The productivity cutoff  $\underline{z}$  is given by:

$$\underline{z}^{\eta} = \frac{\eta \lambda}{\eta - 1} \frac{r^*}{\alpha} \frac{x}{y},\tag{17}$$

while aggregate income y is split between factors as follows:

$$w\ell = (1 - \alpha)y, \qquad r^*\kappa = \alpha \frac{\eta - 1}{\eta}y \qquad and \qquad \Pi = \frac{\alpha}{\eta}y,$$
 (18)

where  $\Pi(t) \equiv \int \pi_t(a, z) d\mathcal{G}_t(a, z)$  is aggregate entrepreneurial income (profits).

<sup>&</sup>lt;sup>14</sup>Specifically,  $\kappa(t) = \int k_t(a, z) d\mathcal{G}_t(a, z)$  and  $\ell(t) = \int n_t(a, z) d\mathcal{G}_t(a, z)$ . Below, aggregate output in (15) equals  $y(t) = \int A(t) (zk_t(a, z))^{\alpha} n_t(a, z)^{1-\alpha} d\mathcal{G}_t(a, z)$ , integrating individual outputs in (4), and expressing it in terms of aggregate capital and labor,  $\kappa(t)$  and  $\ell(t)$ . See derivations in Appendix A1.2.

The first part of Lemma 2 expresses equilibrium aggregates as functions of entrepreneurial wealth x and labor supply  $\ell$ . In contrast to a neoclassical world, entrepreneurial wealth is essential for production in a frictional environment, and it affects aggregate output with elasticity  $\gamma$ . Parameter  $\gamma$  increases in capital intensity  $\alpha$  and in the heterogeneity of entrepreneurs' productivity (decreases in  $\eta$ ), capturing the extent of entrepreneurial rents,  $\Pi = \frac{\alpha}{\eta} y$  in (18). Therefore,  $\gamma$  is a measure of distance from the frictionless limit, and it plays an important role in the analysis of optimal policies in Section 3. Also note that the derived aggregate productivity  $\Theta$  is equally shaped by the primitive productivity A and the financial constraint  $\lambda$ , which together govern endogenous TFP.

Given aggregate production in (16), both the equilibrium wage rate  $w = (1 - \alpha)y/\ell$ and the productivity cutoff  $\underline{z}$  in (17) are increasing functions of  $x/\ell$ . High entrepreneurial wealth x increases capital demand and allows a given labor supply to be absorbed by a smaller subset of more productive entrepreneurs. This raises both the average productivity of active entrepreneurs and aggregate labor productivity (hence wages). If labor supply  $\ell$  increases, less productive entrepreneurs need to become active to absorb it, which in turn reduces average productivity and wages. Note that the tractable functional forms in the expressions of the lemma are due to the Pareto productivity assumption.

The second half of Lemma 2 characterizes the split of aggregate income y in the economy with financial frictions. The share of labor equals  $(1-\alpha)$ , as in the frictionless world, since the choice of labor is unconstrained. However, the presence of financial frictions results in active entrepreneurs making positive profits,  $\Pi > 0$ , in contrast with the neoclassical limit, where  $\Pi = 0$ . Hence, a fraction of national income is received by entrepreneurs at the expense of rentiers, whose share of income,  $r^* \kappa / y$ , falls below  $\alpha$ . This is a result of the depressed demand for capital in a frictional environment, despite the maintained rate of return on capital  $r^*$ . Nonetheless, incomes of all groups in the economy—workers, entrepreneurs and rentiers increase in aggregate output y, which is itself an increasing function of both entrepreneurial wealth x and labor supply  $\ell$ .

Finally, integrating (12) across all entrepreneurs, aggregate entrepreneurial wealth evolves according to:

$$\dot{x} = \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x, \tag{19}$$

where by (18) the first term on the right-hand side equals aggregate entrepreneurial profits  $\Pi$ . Therefore, greater labor supply increases output, which raises entrepreneurial profits and speeds up wealth accumulation, which in future periods leads to higher labor productivity  $y/\ell$  by raising the cutoff productivity level  $\underline{z}$ , according to Lemma 2. A competitive equilibrium in this small open economy is defined in the usual way. Workers and entrepreneurs solve their respective problems taking prices as given, while the path of wages clears the labor market at each point in time and capital is in perfectly elastic supply at interest rate  $r^*$ . Equilibrium can be summarized as a time path for aggregate variables  $\{c, \ell, b, y, x, w, z\}_{t\geq 0}$  satisfying (2), standard household optimality, and (16)–(19), given an initial household asset position  $b_0$ , initial entrepreneurial wealth  $x_0$ , and a path of exogenous productivity A. Actions of individual entrepreneurs can then be recovered from (8)–(10) and (12). This tractable characterization of the transition dynamics in our heterogeneous agent model is what allows for the closed-form optimal policy analysis in Section 3.<sup>15</sup>

### 2.3 Inefficiency: the return wedge

The key to understanding the rationale for policy intervention in our economy is that entrepreneurs earn an excess return relative to workers. Indeed, workers face a rate of  $r^*$ , while an entrepreneur with productivity z generates a return  $R(z) \equiv r^* \left(1 + \lambda \left[\frac{z}{z} - 1\right]^+\right) \ge r^*$ , with  $R(z) > r^*$  for  $z > \underline{z}$ . Because of the collateral constraint, an entrepreneur with productivity  $z > \underline{z}$  cannot expand her capital to drive down her return towards  $r^*$ . Similarly, not only individual entrepreneurs but also entrepreneurs as a group earn an excess return. In particular, the average rate of return across entrepreneurs is given by:

$$\mathbb{E}_z R(z) = r^* \left( 1 + \frac{\lambda \underline{z}^{-\eta}}{\eta - 1} \right) = r^* + \frac{\alpha}{\eta} \frac{y}{x} > r^*, \tag{20}$$

where the first equality integrates R(z) using the Pareto distribution  $G_z(z)$  and the second equality uses the equilibrium cutoff expression (17).

Given that workers and entrepreneurs face different rates of return, which fail to equalize due to the financial friction, a Pareto improvement can be achieved by a wealth transfer from workers to all entrepreneurs (independently of their productivity) combined with a reverse transfer at a later point in time.<sup>16</sup> This perturbation sharply illustrates the nature of the

<sup>&</sup>lt;sup>15</sup>Some of our results could be illustrated in an economy without heterogeneity, with a single productivity type. A model with heterogeneity is, however, closer to the canonical framework used in the macrodevelopment literature, and it allows us to study the effects on misallocation and TFP. Furthermore, and perhaps surprisingly, the presence of a continuum of heterogeneous entrepreneurs gives greater tractability to the model, by summarizing the effects of financial frictions via a single endogenously-evolving productivity cutoff  $\underline{z}$ .

<sup>&</sup>lt;sup>16</sup>More precisely, we show in Appendix A2 that any transfer of resources from workers to entrepreneurs at t = 0 and a reverse transfer at t' > 0 with interest accumulated at a rate  $r_{\omega} = r^* + \omega \int_0^{t'} \frac{\alpha}{\eta} \frac{y(t)}{x(t)} dt > r^*$  for some  $\omega \in (0, 1)$  would necessarily lead to a strict Pareto improvement for all workers and entrepreneurs.

inefficiency in the laissez-faire equilibrium and provides a natural benchmark for thinking about various other policy interventions. Yet such transfers may not be a realistic policy option. For example, large transfers to entrepreneurs may be infeasible for budgetary, distributional or political economy reasons, or due to the associated informational frictions and informational requirements to administer them (see further discussion in Appendix A2.3).

Furthermore, the type of transfer policy discussed here effectively allows the government to get around the specific financial constraint we have adopted, and hence it is not particularly surprising that it results in a Pareto improvement. Such a transfer policy may also not prove robust to alternative formulations of the financial friction. It is for these reasons that the main focus of the paper is on Ramsey-optimal taxation with a given set of simple policy tools. While also having the capacity to Pareto-improve upon the laissez-faire allocation, the policy tools we study in the next section constitute a more realistic and, we think, more robust alternative to transfers.

### **3** Optimal Policy in a One-Sector Economy

In this section we study optimal Ramsey interventions with a given set of policy tools. We start our analysis with two tax instruments—a labor income tax and a savings tax—operating directly on the decision margins of the households. In Section 3.3, we generalize our results to an environment which allows for additional tax instruments directly affecting the decisions of entrepreneurs, including a credit subsidy.

#### 3.1 Economy with taxes

In the presence of labor income and savings taxes on workers,  $\tau_{\ell}(t)$  and  $\tau_{b}(t)$ , the budget constraint of the households changes from (2) to:

$$c + \dot{b} \le (1 - \tau_{\ell})w\ell + (r^* - \tau_b)b + T,$$
(21)

where T are the lump-sum transfers from the government (lump-sum taxes if negative). In our framework, Ricardian equivalence applies, and only the combined wealth of households and the government matters. Therefore, without loss of generality, we assume that the government budget constraint is balanced period by period:

$$T = \tau_\ell w \ell + \tau_b b. \tag{22}$$

More generally, if the government can run a budget deficit and issue debt, we can guarantee

implementation of the Ramsey policies without lump-sum taxes.<sup>17</sup>

In the presence of taxes, the optimality conditions of households become:

$$\dot{u}_c/u_c = \rho - r^* + \tau_b,\tag{23}$$

$$-u_{\ell}/u_{c} = (1 - \tau_{\ell})w, \tag{24}$$

while the wage rate w still satisfies the labor demand relationship (18).

The following result simplifies considerably the analysis of the optimal policies:

**Lemma 3** Any aggregate allocation  $\{c, \ell, b, x\}_{t\geq 0}$  satisfying

$$c + \dot{b} = (1 - \alpha)y(x, \ell) + r^*b,$$
 (25)

$$\dot{x} = \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x, \qquad (26)$$

where  $y(x, \ell)$  is defined in (16), can be supported as a competitive equilibrium under appropriately chosen policies  $\{\tau_{\ell}, \tau_{b}, T\}_{t\geq 0}$ , and the equilibrium characterization in Lemma 2 applies.

Intuitively, equations (25) and (26) are respectively the aggregate budget constraints of workers and entrepreneurs, where we have substituted the government budget constraint (22) and the allocation of aggregate income  $y(x, \ell)$  from Lemma 2. Lemma 2 still holds in this environment because it only relies on labor market clearing and policy functions of the entrepreneurs, which are not affected by the introduced policy instruments. The additional two constraints on the equilibrium allocation are the optimality conditions of workers, (23) and (24), but they can always be ensured to hold by an appropriate choice of labor and savings subsidies for workers,  $\tau_{\ell}$  and  $\tau_{b}$ . Finally, given a dynamic path of  $\ell$  and x, we can recover all remaining aggregate variables supporting the allocation from Lemma 2.

Similarly to the *primal approach* in the Ramsey taxation literature (e.g. Lucas and Stokey, 1983), Lemma 3 allows us to replace the problem of choosing a time path of the policy instruments in a decentralized dynamic economy by a simpler problem of choosing a dynamic aggregate allocation satisfying the *implementability constraints* (25) and (26). These two constraints differ somewhat from those one would obtain following the standard procedure of the primal approach because we exploit the special structure of our model (summarized in Lemma 2) to derive more tractable conditions.

<sup>&</sup>lt;sup>17</sup>As we show below, in the long run  $\tau_b = 0$  and  $\tau_\ell > 0$ , so that the government can roll forward the debt it has accumulated in the short run without violating the intertemporal budget constraint. This can be achieved either without any lump-sum taxes or transfers,  $T \equiv 0$ , or only with lump-sum transfers to households,  $T \geq 0$ , in case the government runs a gross budget surplus in the long run.

#### **3.2** Optimal Ramsey policies

We assume for now that the planner maximizes the welfare of households and puts zero weight on the welfare of entrepreneurs. As will become clear, this is the most conservative benchmark for our results. The Ramsey problem in this case is to choose policies  $\{\tau_{\ell}, \tau_{b}, T\}_{t\geq 0}$ to maximize household utility (1) subject to the resulting allocation being a competitive equilibrium. From Lemma 3, this problem is equivalent to maximizing (1) with respect to the aggregate allocation  $\{c, \ell, b, x\}_{t\geq 0}$  subject to (25)–(26), which we rewrite as:

$$\max_{\{c,\ell,b,x\}_{t\geq 0}} \int_0^\infty e^{-\rho t} u(c,\ell) dt$$
  
subject to  $c+\dot{b} = (1-\alpha)y(x,\ell) + r^*b,$   
 $\dot{x} = \frac{\alpha}{\eta}y(x,\ell) + (r^* - \delta)x,$  (P1)

and given initial conditions  $b_0$  and  $x_0$ . (P1) is a standard optimal control problem with controls  $(c, \ell)$  and states (b, x), and we denote the corresponding co-state vector by  $(\mu, \mu\nu)$ . To ensure the existence of a finite steady state, we assume  $\delta > \rho = r^*$ , which however is *not* essential for the pattern of optimal policies along the transition path.<sup>18</sup>

Before characterizing the solution to (P1), we provide a brief discussion of the nature of this planner's problem. Apart from the Ramsey-problem interpretation that we adopt here, this planner's problem admits two additional interpretations. First, it corresponds to the planner's problem adopted in Caballero and Lorenzoni (2014), which rules out any transfers or direct interventions into the decisions of agents, and only allows for aggregate market interventions which affect agent behavior by moving equilibrium prices (wages in our case). Second, the solution to this planner's problem is a constrained-efficient allocation under the definition developed in Dávila, Hong, Krusell, and Ríos-Rull (2012) for economies with exogenously incomplete markets and borrowing limits, as ours, where standard notions of constrained efficiency are hard to apply. Under this definition, the planner can choose policy functions for all agents respecting, however, their budget sets and exogenous borrowing constraints. Indeed, in our case the planner does not want to change the policy functions of entrepreneurs, but chooses to manipulate the policy functions of households exactly in the way prescribed by the solution to (P1). The implication is that the planner in this case does not need to separate workers and entrepreneurs, relaxing the informational requirement

<sup>&</sup>lt;sup>18</sup>This assumption is not needed if workers are hand-to-mouth in equilibrium or subject to idiosyncratic income risk, in which case  $\delta = \rho > r^*$  is a natural assumption in a small open economy and would also arise endogenously in a closed economy (Aiyagari, 1994).

of the Ramsey policy. As we show in later sections, the baseline structure of the planner's problem (P1) is maintained in a number of extensions we consider.

The optimality conditions for the planner's problem (P1) are given by (see Appendix A3.1):

$$\frac{\dot{u}_c}{u_c} = \rho - r^* = 0,$$
(27)

$$-\frac{u_{\ell}}{u_c} = \left(1 - \gamma + \gamma\nu\right) \cdot (1 - \alpha)\frac{y}{\ell},\tag{28}$$

$$\dot{\nu} = \delta\nu - \left(1 - \gamma + \gamma\nu\right)\frac{\alpha}{\eta}\frac{y}{x}.$$
(29)

An immediate implication of (27) is that the planner does not distort the intertemporal margin, that is  $\tau_b \equiv 0$ , as follows from (23). There is no need to distort the workers' saving decision since, holding labor supply constant, consumption does not have a direct effect on output  $y(x, \ell)$  and hence on wealth accumulation in (26).<sup>19</sup>

In contrast, the laissez-faire allocation of labor, which satisfies (24) with  $\tau_{\ell} \equiv 0$ , is in general suboptimal. Indeed, combining the planner's optimality condition (28) with (18) and (24), the *labor wedge* (tax) can be expressed in terms of the co-state  $\nu \geq 0$  as:

$$\tau_{\ell} = \gamma (1 - \nu), \tag{30}$$

and whether labor supply is subsidized or taxed depends on whether  $\nu$  is greater than one.

Indeed, the planner has two reasons to distort the choice of labor supply,  $\ell$ . First, workers take wages as given and do not internalize that  $w = (1 - \alpha)y/\ell$  (see Lemma 2); that is, by restricting labor supply workers can increase their wages. As the planner only cares about the welfare of workers, this *monopoly effect* induces the planner to reduce labor supply. The statically optimal monopolistic labor tax equals  $\gamma$  in our model, and corresponds to the first term in brackets in (30).

Second, workers do not internalize the positive effect of their labor supply on entrepreneurial profits and wealth accumulation, which affects future output and wages. This *dynamic productivity effect* through wealth accumulation forces the planner to increase labor supply, and it is reflected in the second term in (30),  $-\gamma\nu$ . When entrepreneurial wealth is scarce, its shadow value for the planner is high ( $\nu > 1$ ), and the planner increases labor supply,  $\tau_{\ell} < 0.^{20}$ Otherwise, the static consideration dominates, and the planner reduces (taxes) labor supply.

<sup>&</sup>lt;sup>19</sup>In a closed economy, in addition to intervening in the labor market, the planner also chooses to distort the intertemporal savings margin to encourage a faster accumulation of capital (see Appendix A4.3).

<sup>&</sup>lt;sup>20</sup>Solving (29) forward,  $\nu$  can be expressed as a net present value of future marginal products of wealth,  $\partial y/\partial x = \gamma y/x$ , which are monotonically decreasing in x, with  $\lim_{x\to 0} \partial y/\partial x = \infty$  (see Appendix A3.2).



Figure 1: Planner's allocation: phase diagram for transition dynamics

Finally, recall that  $\gamma$  is a measure of the distortion arising from the financial frictions, and in the frictionless limit with  $\gamma \to 0$  the planner does not need to distort any margin.

The planner's optimal allocation  $\{c, \ell, b, x\}_{t\geq 0}$  solves the dynamic system (25)–(29). With  $r^* = \rho$ , the marginal utility of consumption is constant over time,  $u_c(t) \equiv \bar{\mu}$ , and the system separates in a convenient way. Given a level of  $\bar{\mu}$ , which can be pinned down from the intertemporal budget constraint, the optimal labor wedge  $\tau_{\ell} = \gamma(1 - \nu)$  can be characterized by means of two ODEs in  $(x, \nu)$ , (26) and (29), together with the static optimality condition (28). These can be analyzed by means of a phase diagram (Figure 1) and other standard tools (see Appendix A3.2) to yield:

**Proposition 1** The solution to the planner's problem (P1) corresponds to the saddle path of the ODE system (26) and (29), as summarized in Figure 1. In particular, starting from  $x_0 < \bar{x}$ , both x(t) and  $\tau_{\ell}(t) = \gamma (1 - \nu(t))$  increase over time towards the unique positive and globally stable steady state  $(\bar{\tau}_{\ell}, \bar{x})$ , with labor supply taxed in steady state:

$$\bar{\tau}_{\ell} = \frac{\gamma}{\gamma + (1 - \gamma)(\delta/\rho)} > 0.$$
(31)

Labor supply is subsidized,  $\tau_{\ell}(t) < 0$ , when entrepreneurial wealth x(t) is low enough. The planner does not distort the workers' intertemporal margin,  $\tau_{b}(t) \equiv 0$ .

The optimal steady-state labor wedge is strictly positive, meaning that in the long run the planner suppresses labor supply rather than subsidizing it. This tax is however smaller than the optimal monopoly tax equal to  $\gamma$  (i.e.,  $0 < \bar{\tau}_{\ell} < \gamma$ ), because with  $\delta > r^*$  entrepreneurial wealth accumulation is bounded and the financial friction is never resolved (i.e., even in steady state the shadow value of entrepreneurial wealth is positive,  $\bar{\nu} > 0$ ). Nonetheless, in steady state the redistributive force necessarily dominates dynamic productivity considerations. This, however, is not the case along the entire transition path, as we prove in Proposition 1 and illustrate in Figure 1. Consider a country that starts out with entrepreneurial wealth considerably below its steady-state level, i.e. in which entrepreneurs are initially severely undercapitalized. Such a country finds it optimal to increase (subsidize) labor supply during the initial transition phase, until entrepreneurial wealth reaches a high enough level.

Figures 2 and 3 illustrate the transition dynamics for key variables, comparing the allocation chosen by the planner to the one that would obtain in a laissez-faire equilibrium.<sup>21</sup> The left panel of Figure 2 plots the optimal labor tax, which is negative in the early phase of the transition (i.e., a labor supply subsidy), and then switches to being positive in the long run. This is reflected in the initially increased and eventually depressed labor supply in the planner's allocation in Figure 3a. The purpose of the labor supply subsidy is to speed up entrepreneurial wealth accumulation (Figures 2b and 3b), which in turn translates into higher productivity and wages in the medium run, at the cost of their reduction in the short run (Figures 3c and 3d). The labor tax and suppressed labor supply in the long run are used to redistribute the welfare gains from entrepreneurs towards workers through the resulting increase in wages.<sup>22</sup>

Figure 3e shows that during the initial phase of the transition, the optimal policy increases GDP as well as the incomes of all groups of agents—workers, active entrepreneurs and rentiers (inactive entrepreneurs)—according to Lemma 2. Output y is higher both due to a higher labor supply  $\ell$  and increased capital demand  $\kappa$ , while the capital-output ratio  $\kappa/y$  remains constant according to (18). This increase in demand is met by an inflow of capital, which is in perfectly elastic supply in a small open economy. The effect of the increase in inputs  $\ell$  and  $\kappa$  is partly offset by a reduction in TFP due to a lower productivity cutoff  $\underline{z}$ , as less productive entrepreneurs need to become active to absorb the increased labor supply.

<sup>&</sup>lt;sup>21</sup>Our numerical examples use the following benchmark parameter values:  $\alpha = 1/3$ ,  $\delta = 0.1$ ,  $\rho = 0.03$ ,  $\eta = 1.06$  and  $\lambda = 2$ , as well as balanced growth preferences with a constant Frisch elasticity  $1/\varphi$ :  $u(c, \ell) = \log c - \psi \ell^{1+\varphi}/(1+\varphi)$ , with  $\psi = 1$  and  $\varphi = 1$ . The initial condition  $x_0$  is 10% of the steady-state level in the laissez-faire equilibrium, and the initial wealth of workers is  $b_0 = 0$ .

<sup>&</sup>lt;sup>22</sup>Interestingly, even if the reversal of the labor subsidy were ruled out (by imposing a restriction  $\tau_{\ell} \leq 0$ ), the planner still wants to subsidize labor during the early transition, emphasizing that the purpose of this policy is not merely a reverse redistribution at a later date. The same is true in an alternative model where financially-constrained firms are collectively owned by workers, and hence there is no distributional conflict.



Figure 2: Planner's allocation: labor tax  $\tau_{\ell}(t)$  and entrepreneurial wealth x(t)Note: In panel (b), the steady-state entrepreneurial wealth in the laissez-faire equilibrium is normalized to 1.

Although our numerical example is primarily illustrative, it can be seen that the transition dynamics in this economy, where heterogeneous producers face collateral constraints, may take a very long time, consistent with the observed post-war growth miracles (for further discussion see Buera and Shin, 2013). Furthermore, the quantitative effects of the Ramsey labor market policies may be quite pronounced. In particular, in our example the Ramsey policy increases labor supply by up to 18% and GDP by up to 12% during the initial phase of the transition, which lasts around 20 years. This is supported by an initial labor supply subsidy of over 20%, which switches to a 12% labor tax in the long run. Despite the increased labor income, workers initially suffer in flow utility terms (Figure 3f) due to increased labor supply. Workers are compensated with a higher utility in the future, reaping the benefits of both higher wages and lower labor supply, and gain on net intertemporally. We revisit these results in Section 3.4, where we consider overlapping generations of finitely-lived households that also face borrowing constraints (in particular, *cf.* Appendix Figure A8).

**Implementation** The Ramsey-optimal allocation can be implemented in a number of different ways. For concreteness, we focus here on the early phase of the transition, when the planner wants to increase labor supply. The way we set up the problem, the optimal allocation during this initial phase is implemented with a labor supply subsidy,  $\varsigma_{\ell}(t) \equiv -\tau_{\ell}(t) > 0$ , financed by a lump-sum tax on workers (or government debt accumulation). In this case, workers' gross labor income including subsidy is  $(1 + \varsigma_{\ell})(1 - \alpha)y$ , while their net income subtracting the lump-sum tax is still given by  $(1 - \alpha)y$ , hence resulting in no direct change



Figure 3: Planner's allocation: proportional deviations from the laissez-faire equilibrium

Note: In panel (d), the deviations in TFP are the same as the deviations in  $\underline{z}^{\alpha}$ , as follows from (15). In panel (e), income deviations characterize simultaneously the deviations in output (y), wage bill  $(w\ell)$ , profits ( $\Pi$ ), capital income  $(r^*\kappa)$ , and hence capital  $(\kappa)$ , as follows from Lemma 2.

in their budget set. Note that increasing labor supply unambiguously increases net labor income  $(w\ell)$ , but decreases the net wage rate (w) paid by firms. This is why we sometimes refer to this policy as *wage suppression*.

An equivalent implementation is to give a wagebill subsidy to firms financed by a lumpsum tax on workers. In this case, the equilibrium wage rate increases, but the firms pay only a fraction of the wage bill, and the resulting allocation is the same. There are of course alternative implementations that rely on directly controlling the *quantity* of labor supplied, rather than its *price*, such as *forced labor*—a forced increase in the hours worked relative to the competitive equilibrium. Such a non-market implementation pushes workers off their labor supply schedule and the wage is determined by moving along the labor demand schedule of firms. Our theory is silent on the relative desirability of one form of intervention over another. See Weitzman (1974) for a discussion. Furthermore, desirable allocations may be achieved without any tax interventions by means of market regulation, e.g. by shifting the bargaining power from workers to firms in the labor market, as is often the case in practice (see Online Appendix B and Appendix A2.3 for further discussion).

The general feature of all these implementation strategies is that they make workers work hard even though wages paid by firms are low. Put differently, the common feature of all policies is their *pro-business* tilt in the sense that they reduce the effective labor costs to firms, allowing them to expand production and generate higher profits, in order to facilitate the accumulation of wealth in the absence of direct transfers to entrepreneurs.

Learning-by-doing analogy One alternative way of looking at the planner's problem (P1) is to note that from (16), GDP depends on current labor supply  $\ell(t)$  and entrepreneurial wealth x(t). From (19), entrepreneurial wealth accumulates as a function of past profits, which are a constant fraction of past aggregate incomes, or outputs. Therefore, current output depends on the entire history of past labor supplies,  $\{\ell\}_{t\geq 0}$ , and the initial level of wealth,  $x_0$ . Importantly, in the competitive equilibrium workers do not take into account the effect of their labor supply decisions on the accumulation of this state variable. In contrast, the planner internalizes it. This setup, hence, is isomorphic at the aggregate to a model of a small open economy with a learning-by-doing externality in production (see, for example, Krugman, 1987; Matsuyama, 1992). Entrepreneurial wealth in our setup plays the same role as physical productivity in theories with learning-by-doing. As a result, some of our policy implications have a lot in common with those that emerge in economies with learning-by-doing externalities, as we discuss in Section 5. That being said, the detailed micro-structure of our environment not only provides discipline for the aggregate planning problem, but

also differs in qualitative ways from an environment with learning-by-doing. For example, as explained above (and in more detail in Appendix A2.2), transfers between entrepreneurs and workers would be a powerful tool in our environment, but have no bite in an economy with learning-by-doing.

**Pareto weight on entrepreneurs** Our analysis generalizes in a natural way to the case where the planner puts an arbitrary non-zero Pareto weight on the welfare of the entrepreneurs. In Appendix A1.3, we derive the expected present value of an entrepreneur with assets  $a_0$  at time t = 0, denoted  $V_0(a_0)$ . In Appendix A4.2 we extend the baseline planner's problem (P1) to allow for an arbitrary Pareto weight,  $\theta \ge 0$ , on the utilitarian welfare function for all entrepreneurs,  $\mathbb{V}_0 = \int V_0(a) dG_{a,0}(a)$ . We show that the resulting optimal policy parallels that characterized in our main Proposition 1, with the optimal labor tax now given by:

$$\tau_{\ell}^{\theta}(t) = \gamma \left(1 - \nu(t)\right) - \theta \gamma \frac{e^{(\rho - \delta)t}}{\delta \bar{\mu} x(t)}.$$
(32)

Therefore, the optimal tax schedule simply shifts down (for a given value of  $\nu$ ) in response to a greater weight on the entrepreneurs in the social objective. That is, the transition is associated with a larger subsidy to labor supply initially and a smaller tax on labor later on. In this sense, we view our results above as a conservative benchmark, since even when the planner does not care about entrepreneurs, she still chooses a pro-business policy tilt during the early transition.

#### **3.3** Additional tax instruments

In order to evaluate the robustness of our conclusions, we now briefly consider the case with additional tax instruments which directly affect the decisions of entrepreneurs. In particular we introduce a capital subsidy  $\varsigma_k$ , which in our environment is equivalent to a credit subsidy.<sup>23</sup> The key result of this section is that, despite the availability of this more direct policy instrument to address financial constraints, it is nevertheless optimal to distort workers' labor supply decisions by suppressing wages early on during the transition and increasing them in the long run. In Appendix A3.3, we characterize a more general case, which additionally allows for a revenue (sales) subsidy, a profit subsidy, and an asset subsidy to entrepreneurs.

Specifically, we now consider the profit maximization of an entrepreneur that faces a

<sup>&</sup>lt;sup>23</sup>Indeed, a subsidy to  $r^*k$  in our model is equivalent to a subsidy to  $r^*(k-a)$ , as all active entrepreneurs choose the same leverage,  $k = \lambda a$ .

wagebill subsidy  $\varsigma_w$  and a cost of capital subsidy  $\varsigma_k$ :

$$\pi(a,z) = \max_{\substack{n \ge 0, \\ 0 \le k \le \lambda a}} \left\{ A(zk)^{\alpha} n^{1-\alpha} - (1-\varsigma_w) wn - (1-\varsigma_k) r^* k \right\}.$$
(33)

Credit (capital) subsidies are, arguably, a natural tax instrument to address the financial friction, and they have been an important element of real world industrial policies (see Online Appendix B as well as McKinnon, 1981; Diaz-Alejandro, 1985; Leipziger, 1997).

In the presence of the additional subsidies to entrepreneurs, the equilibrium characterization in Lemma 2 no longer applies and needs to be generalized, as we do in Appendix A3.3. In particular, we show that the aggregate output function now generalizes (16) and is given by:

$$y(x,\ell) = (1-\varsigma_k)^{-\gamma(\eta-1)} \Theta x^{\gamma} \ell^{1-\gamma},$$

with  $\gamma$  and  $\Theta$  defined as before. Furthermore, the planner's problem has a similar structure to (P1), with the added optimization over the choice of the additional subsidies. This allows us to prove the following:

**Proposition 2** When the planner's only policy tools are a wagebill subsidy and a capital subsidy to entrepreneurs, the optimal Ramsey policy is to use both of them in tandem, and set them according to:

$$\frac{\varsigma_w}{1-\varsigma_w} = \frac{\varsigma_k}{1-\varsigma_k} = \frac{\alpha}{\eta}(\nu-1),\tag{34}$$

where  $\nu$  is the shadow value of entrepreneurial wealth, which evolves as described in Section 3.2.

The key implication of Proposition 2 is that even when a credit (capital) subsidy  $\varsigma_k$  is available, the planner still finds it optimal to use the labor (wagebill) subsidy  $\varsigma_w$  alongside it. This is because credit subsidies introduce distortions of their own by affecting the extensive margin of selection into entrepreneurship.<sup>24</sup> As a result, the planner prefers to combine both instruments in order to minimize the amount of created deadweight loss. Furthermore, note that the two subsidies are perfectly coordinated, leaving undistorted the capital-labor ratio chosen by the entrepreneurs. Lastly, observe that the shadow value of entrepreneurial wealth  $\nu$  is, as before, a sufficient state variable for the stance of the optimal policy, given the

<sup>&</sup>lt;sup>24</sup>Note that the most direct way to address the financial friction is to relax the collateral constraint (5) by increasing  $\lambda$ , which would lead in equilibrium to reallocation of capital from less to more productive entrepreneurs and exit of the marginal ones. In contrast, a capital subsidy leads to additional entry on the margin, resulting in greater production inefficiency and lower TFP.

parameters of the model.<sup>25</sup> When  $\nu > 1$ , entrepreneurial wealth is scarce, and the planner subsidizes both entrepreneurial production margins. As wealth accumulates,  $\nu$  declines and eventually becomes less than 1, a point at which the planner starts taxing both margins, just like in Proposition 1. As a general principle, whenever entrepreneurial wealth is scarce, the planner utilizes all available policy instruments in a pro-business manner.

Lastly, we briefly comment on wealth transfers as a policy tool. In our analysis, we ruled out direct redistribution of wealth, either between entrepreneurs of different productivities, or between entrepreneurs and workers. In Appendix A2.2, we relax the latter restriction and allow for direct transfers between workers and entrepreneurs, which in certain cases can also be engineered using a set of available distortionary taxes (see Appendix A3.3). Here again our conclusion regarding the optimality of a labor subsidy when entrepreneurial wealth is low remains intact, as long as the feasible transfers are finite. Put differently, the only case in which there is no benefit from increasing labor supply in the initial transition phase is when an unbounded transfer from workers to entrepreneurs is available, which allows the planner to immediately jump the economy to its steady state.<sup>26</sup>

#### 3.4 Finite lives and household borrowing constraints

We now extend our analysis to overlapping generations (OLG) of workers, who face a hazard of dying and are replaced by new generations, as in Blanchard (1985) and Yaari (1965). Later, we additionally introduce borrowing constraints on the workers, to capture the idea that financial frictions directly affect all agents in the economy. One may suspect that the planner's priorities and allocations change considerably in these cases, because now the costs and benefits of the policies are distributed unevenly across generations. Yet, we show that our main insights are robust, perhaps surprisingly, to these extensions.

We assume that a worker lives to age  $s \ge 0$  with probability  $e^{-qs}$ , where q is an instantaneous death rate common across all age groups. The results can be extended to the case of non-constant death hazard over the life cycle as in Calvo and Obstfeld (1988). At each time t, agents are born at the same rate q so that the total population is stable, and normalized to one. Hence, at any point in time, the number (density) of s-year olds is  $qe^{-qs}$ . Individuals born at date  $\tau$  (cohort  $\tau$ ) have lifetime utility  $U(\tau) = \int_{\tau}^{\infty} e^{-(\rho+q)(t-\tau)}u_{\tau}(t)dt$ ,

<sup>&</sup>lt;sup>25</sup>Compare (34) with (30): in both cases the optimal subsidies are proportional to  $\frac{\alpha}{\eta}(\nu-1)$ , given the definition of  $\gamma$  in (16). Also note that the same allocation as in Proposition 2 can be achieved by replacing the wage subsidy  $\varsigma_w$  with a labor income subsidy, setting  $\tau_{\ell} = -\frac{\varsigma_w}{1-\varsigma_w} = \frac{\alpha}{\eta}(1-\nu)$ .

<sup>&</sup>lt;sup>26</sup>With unlimited transfers, the planner can fully relax the aggregate financial constraint of entrepreneurs in (P1), and hence ensure  $\nu \equiv 1$  in every period and avoid the need to use distortionary policy instruments.

where  $u_{\tau}(t) = u(c_{\tau}(t), \ell_{\tau}(t))$  is the period utility at time t of a member of cohort  $\tau$ . We further assume that the wealth of the dying workers is passed on to the surviving generations of workers, via bequest or a perfect annuity market. Then, aggregating over the cross-sectional age distribution, the resulting budget constraint of the household sector is still given by (25). Since nothing changes on the side of entrepreneurs, the planner faces the same implementability conditions (25)–(26), and Lemma 3 still applies.

It remains to specify the planner's objective. In particular, we need to take a stand on how the planner weighs cohorts born at different dates. We assume that the planner discounts the lifetime utilities of different generations at rate  $\rho$ , a rate which need not equal the individual time preference rate  $\rho$ . We further follow Calvo and Obstfeld (1988) and assume that social welfare evaluated at date 0 is given by:

$$W_{0} = \int_{-\infty}^{\infty} e^{-\rho\tau} q \, U_{0}(\tau) \mathrm{d}\tau, \qquad \text{where} \qquad U_{0}(\tau) = \begin{cases} U(\tau), & \tau \ge 0, \\ \int_{0}^{\infty} e^{-(\rho+q)(t-\tau)} u_{\tau}(t) \mathrm{d}t, & \tau < 0. \end{cases}$$
(35)

In words,  $U_0(\tau)$  is the remaining lifetime utility as of date 0 for cohort  $\tau$ , discounted to the date of birth.<sup>27</sup> That is, the planner maximizes  $W_0$ , which uses her time preference rate  $\rho$  to aggregate  $U_0(\tau)$  for all cohorts  $\tau \in (-\infty, \infty)$ , and where q is each cohort's size at birth.

Next, using a change of variable from cohort  $\tau$  to age  $s = t - \tau$ , we rewrite the welfare criterion in (35) as:

$$W_0 = \int_0^\infty e^{-\varrho t} V(t) \mathrm{d}t, \qquad V(t) \equiv \int_0^\infty q e^{-qs} \cdot e^{-(\rho-\varrho)s} \cdot u\big(\tilde{c}(t,s), \tilde{\ell}(t,s)\big) \mathrm{d}s, \qquad (36)$$

where  $\tilde{c}(t,s)$  and  $\ell(t,s)$  are the consumption and labor supply of *s*-year-old workers at time *t*. Intuitively, V(t) represents the utility flow from all workers alive at time *t*, aggregating across the cross-sectional age distribution with density  $qe^{-qs}$ , with  $(\rho - \varrho)$  reflecting the relative weight the planner puts on younger generations at a given point in time. The key insight of Calvo and Obstfeld (1988) is that optimal allocations can be conveniently found by means of a two-step procedure. First, statically maximize V(t) subject to the constraint that the integrals of  $\tilde{c}(t,s)$  and  $\tilde{\ell}(t,s)$  equal aggregate consumption and labor supply c(t) and  $\ell(t)$ : formally,  $\int_0^{\infty} qe^{-qs}\tilde{c}(t,s)ds \leq c(t)$  and  $\int_0^{\infty} qe^{-qs}\tilde{\ell}(t,s)ds \geq \ell(t)$ . Second, choose the time path of c(t) and  $\ell(t)$  that maximizes  $W_0$ .

<sup>&</sup>lt;sup>27</sup>It may seem somewhat unnatural to discount the utility of those already alive back to their birthdates  $\tau \leq 0$ . However, Calvo and Obstfeld (1988) show that this approach (unlike others) results in a time-consistent planner's objective, even when  $\rho \neq \rho$  (see Appendix A3.4).

We first consider a benchmark case of  $\rho = \rho$ , that is of a planner who places equal weight on all generations. It is then intuitive that the planner gives the same allocation to people of all ages at any given point in time,  $\tilde{c}(t,s) \equiv c(t)$  and  $\tilde{\ell}(t,s) \equiv \ell(t)$  for all s. Therefore, the planner's objective in (36) simply becomes  $W_0 = \int_0^\infty e^{-\rho t} u(c(t), \ell(t)) dt$ , equivalent to that in (P1). That is, when  $\rho = \rho$ , the planner's problem with finitely-lived workers is completely isomorphic to the case with infinitely-lived workers, and as a result *none* of the optimal policy implications change in any way relative to Proposition 1 and Figures 1–3. While this result may seem surprising at first, it is nonetheless intuitive: when a worker dies she is replaced by another worker with an identical utility flow from allocations, and since the planner puts equal weight on these two workers, her objective is unaffected by finite lives.<sup>28</sup>

Next consider the case with  $\rho > \rho$  capturing the planner's solidarity with earlier generations. From (36), one can then see that an unconstrained planner would discriminate between older and younger generations by allocating less consumption to younger workers at a given point in time. Such allocations are arguably unnatural because implementing them requires age-dependent tax instruments, and therefore we impose an additional constraint on the planner that  $\tilde{c}(t,s) \equiv c(t)$  and  $\tilde{\ell}(t,s) \equiv \ell(t)$  for all s at any given point in time t. As a result,  $V_t = \frac{q}{\rho+q-\varrho}u(c(t),\ell(t))$ , and the planner's objective is  $W_0 \propto \int_0^\infty e^{-\varrho t}u(c(t),\ell(t))dt$ . Therefore, the analysis is still isomorphic to solving the problem in (P1), but now with a higher discount rate  $\rho > \rho = r^*$ . A natural upper limit on the planner's discount rate is  $\rho = \rho + q$ , which is equivalent to the planner giving an exclusive weight to the earliest cohort.<sup>29</sup>

In Appendix A3.4, we show that all optimality conditions in this case are unchanged, except for (27) which becomes  $\dot{u}_c/u_c = \rho - r^* > 0$ , i.e. the planner chooses to front-load consumption. The characterization of the optimal policy (29)–(30), however, is unchanged, and the qualitative pattern of the initially increased labor supply and lowered wages still applies. In fact, if utility features no income effects (i.e., under GHH preferences), the time path of the optimal labor tax  $\tau_{\ell}(t)$  is independent of the value of  $\rho$ . Therefore, the main insights of our analysis are robust, perhaps surprisingly, to overlapping generations of workers even under a present-biased planner. Indeed, when the planner can freely borrow in international capital markets, she favors early generations solely via increased consumption,

<sup>&</sup>lt;sup>28</sup>Finite lives, however, require a brief discussion of decentralization of the optimal Ramsey plan. In one case, the new generations of workers need to be endowed with the same wealth as all surviving workers, by means of bequests or government transfers. This is, however, not necessary in the presence of perfect borrowing markets, in which case the government needs to subsidize the consumption of earlier cohorts (when productivity and output is low) by accumulating debt and levying taxes in the long run. We discuss below the case with borrowing constraints, where such transfers are not possible.

<sup>&</sup>lt;sup>29</sup>To see this assume that the planner only cares about the oldest cohort which amounts to maximizing  $\int_0^\infty e^{-(\rho+q)t} u(c(t), \ell(t)) dt$ , corresponding to  $W_0$  in the text when  $\rho = \rho + q$ .

keeping unchanged the optimal policy on the supply side.

This analysis naturally leads to the question of borrowing constraints on households and the planner. First, we are interested whether finite horizons have greater bite in the presence of borrowing constraints. Second, we generalize our analysis to feature financial (borrowing) constraints on all agents in the economy, not just the entrepreneurs. For simplicity, we consider the case in which households cannot borrow at all and are hand-to-mouth, and we assume the planner needs to honor this constraint. We show in Appendix A3.4 that, while the expression for the optimal tax (30) does not change in this case, the shadow value of entrepreneurial wealth  $\nu$  (equation (29)) is different. In particular, the optimal time path of the labor tax becomes less steep, featuring smaller subsidies in the short run. This difference from our baseline is more pronounced the more present-biased the planner is, i.e. the larger is  $\rho - \rho > 0$ . Indeed, without the ability to shift consumption towards earlier generations, the usefulness of the labor wedge is reduced, since it delivers only delayed productivity gains. Nevertheless, it remains true even in this case that low financial wealth of entrepreneurs provides a rationale for a stage-dependent policy intervention that subsidizes labor supply early on, and taxes it later in the transition. We illustrate these results in Appendix Figure A8.

## 4 Quantitative Exploration

In the previous sections we developed a tractable model of transition dynamics with financial constraints, which allows for a sharp analytical characterization of the optimal dynamic policy interventions (see Propositions 1 and 2). Towards this goal, we adopted a number of assumptions which allow for tractable aggregation of the economy and result in a simple characterization of equilibrium dynamics under various government policies (see Lemma 3). This naturally raises the question of robustness of the results to relaxation of the main assumptions, which is one of the goals of this section. Doing so requires giving up analytical tractability and extending the model to a richer quantitative environment. We follow the benchmark quantitative framework in the macro-development literature and calibrate our quantitative model to a typical developing economy. The second goal of this section is to evaluate the *quantitative importance* of alternative policies, not necessarily optimal ones, for welfare and growth in the emerging economies. As we will see, the results in this section confirm our main message that pro-business policies are especially important for growth at earlier stages of development, and that such policies can be welfare-improving even from workers' perspective.

#### 4.1 Quantitative Model

The economy is similar to the baseline model in Section 2 with four main differences. First, production functions now feature decreasing returns to scale. This relaxes the feature of the baseline model that all active producers are collateral-constrained, allowing some of them to grow out of the financial frictions over time. Second, we relax the assumption that productivity shocks are *iid* over time and consider a persistent productivity process. Third and relatedly, in the baseline model the cross-sectional productivity distribution is Pareto with unbounded support. In the model of this section, the productivity process instead evolves on a compact interval and hence the stationary productivity distribution has bounded support. Finally, following the extension in Section 3.4, we assume that not only entrepreneurs but also households are financially constrained.

The entrepreneurs still maximize (3) subject to (6) and collateral constraint (5). However, their production function now features decreasing returns,  $\beta < 1$ :

$$y = A \left[ (zk)^{\alpha} n^{1-\alpha} \right]^{\beta}$$

instead of the constant returns technology in (4). Productivity z follows a jump-diffusion process in logs (described more formally in Appendix A5):

$$d\log z_t = -\nu \log z_t dt + \sigma dW_t + dJ_t.$$
(37)

In the absence of jumps  $(dJ_t = 0)$ , this is an Ornstein-Uhlenbeck process, a continuoustime analogue of an AR(1) process, with mean-reversion  $\nu$  and innovation dispersion  $\sigma$ . We further assume that the process is reflected both above and below, and therefore lives on a bounded interval  $[\underline{z}, \overline{z}]$ . Finally, jumps arrive infrequently at a Poisson rate  $\phi$ , and conditional on a jump a new productivity z' is drawn from a truncated Pareto distribution with tail parameter  $\eta > 1$  and support  $[z, \overline{z}]$ .

In contrast to our baseline model, we now assume that workers cannot borrow, and therefore are hand-to-mouth along the equilibrium growth trajectory. Therefore, they effectively maximize  $u(c, \ell)$  period-by-period subject to  $c = (1 - \tau_{\ell})w\ell + T$ , with the lump-sum transfers distributing the collected tax revenue back to the households. As in the quantitative examples in our baseline model, workers have balanced growth preferences with a constant Frisch elasticity,  $u(c, \ell) = \log c - \frac{1}{1+\varphi} \ell^{1+\varphi}$ . Finally, as in Midrigan and Xu (2014), we assume that an exogenous fraction  $(1 - \omega)$  of the population are workers and a fraction  $\omega$  are entrepreneurs. These changes to the model, particularly the adoption of decreasing returns, imply that the state space of the model now necessarily includes the time-varying joint distribution of endogenous wealth and exogenous productivity,  $\mathcal{G}_t(a, z)$ . The enormous (infinite-dimensional) state space of our quantitative model makes it extremely difficult to analyze optimal policy outside of stationary equilibria. Precisely this problem has been the main impediment to this type of analysis in the earlier literature. In particular, it becomes computationally infeasible to study fully general time-varying optimal policy in which the tax instruments are arbitrary functions of time as is the case in our baseline analysis.<sup>30</sup>

To make progress under these circumstances, we adopt a pragmatic approach and restrict the time-dependence of the policy instruments in a parametric way. Motivated by the results of Section 3, as illustrated in Figures 1 and 2, we restrict the time paths of the tax policy to be an exponential function of time:

$$\tau_{\ell}(t) = e^{-\gamma_{\ell}t} \cdot \underline{\tau}_{\ell} + \left(1 - e^{-\gamma_{\ell}t}\right) \cdot \bar{\tau}_{\ell},\tag{38}$$

parameterized by a triplet of the initial tax rate  $\underline{\tau}_{\ell}$ , the steady-state tax rate  $\bar{\tau}_{\ell}$ , and the convergence rate  $\gamma_{\ell}$ . Under this parameterization, the *half-life* of the policy, or the time it takes to go half way from  $\underline{\tau}_{\ell}$  to  $\bar{\tau}_{\ell}$ , is equal to  $\log 2/\gamma_{\ell}$ . This parametric approach reduces the infinite-dimensional optimal policy problem to one of finding three optimal policy parameters  $(\underline{\tau}_{\ell}, \bar{\tau}_{\ell}, \gamma_{\ell})$ .

Lastly, we assume that the planner chooses these tax parameters to maximize a weighted average of initial welfare of workers and entrepreneurs:

$$\mathcal{V}_0 = (1-\omega) \int_0^\infty e^{-\rho t} u(c_t, \ell_t) \mathrm{d}t + \theta \omega \int v_0(a, z) \mathrm{d}\mathcal{G}_0(a, z).$$
(39)

where  $v_0(a, z)$  is the expected life-time utility of an entrepreneur starting at time t = 0with wealth a and productivity z,  $\omega$  is the population share of entrepreneurs, and  $\theta$  is their Pareto weight in the planner's problem. Appendix A5 spells out in more detail the model's equilibrium conditions, including the system of coupled Hamilton-Jacobi-Bellman and Kolmogorov Forward equations that describe the problem of entrepreneurs and the evolution of the distribution  $\mathcal{G}_t(a, z)$  (see also Achdou, Han, Lasry, Lions, and Moll, 2017).

<sup>&</sup>lt;sup>30</sup>Some existing analyses of optimal policy do take into account transition dynamics but restrict tax instruments to be constant over time, making the optimal policy choice effectively a static problem (see e.g. Conesa, Kitao, and Krueger, 2009). Our main result that the sign of the optimal policy differs depending on the stage of development emphasizes the importance of examining time-varying optimal policy. Some recent work has developed numerical methods for finding social optima with fully time-varying tax instruments (Nuño and Moll, 2018), but this is currently only feasible in simpler environments.

#### 4.2 Parameterization

We parameterize the model to capture relevant features of a typical emerging economy, with an initial condition aimed to represent an early stage of development. Our model is similar to the benchmark quantitative models in the literature, namely Buera and Shin (2013) and Midrigan and Xu (2014), and therefore we follow a similar calibration strategy.<sup>31</sup> In Appendix Table A2, we describe the calibrated parameter values, and we discuss the important ones here, relegating the details to Appendix A5.

First, we take as the initial wealth distribution  $\mathcal{G}_0(a, z)$  a ten-fold scaled-down version of the stationary long-run wealth distribution in the absence of policy. In other words, the initial wealth is one tenth of the final wealth under laissez-faire, while the correlation between a and z is the same. This ten-fold increase in the wealth of entrepreneurs contributes to more than doubling of the GDP along the transition path with growth rates exceeding 5% over the first 12 years of transition.<sup>32</sup> Second, we set the parameter governing the tightness of financial constraints (5) to  $\lambda = 2$ . This results in a steady-state external finance to GDP ratio of 2.3 which is in between the values of the 2011 external finance to GDP ratios of China (2.0) and South Korea (2.5) based on data by Beck, Demirguc-Kunt, and Levine (2000).

Third, the literature on the macroeconomic effects of financial frictions in developing countries emphasizes the importance of the stochastic process for productivity z (Midrigan and Xu, 2014; Moll, 2014). Asker, Collard-Wexler, and de Loecker (2014; henceforth, ACD) have estimated productivity processes for 33 developing countries and we use their estimates to discipline the process in our model. We set  $\nu$  so that the annual autocorrelation of productivity equals 0.85, the average of the country-specific estimates in ACD. We set  $\sigma = 0.3$ , which is towards the lower end of ACD estimates. Lastly, we set  $\phi = 0.1$  implying that Poisson jumps arrive infrequently, on average every ten years.

Finally, we set the population share of entrepreneurs  $\omega$  equal to one third, a high incidence common to developing countries, and considerably higher than in developed countries like the U.S., where it is 10–15% (see e.g. Cagetti and De Nardi, 2006). We experiment with three different values of the Pareto weight on entrepreneurs  $\theta \in \{0, 1/2, 1\}$ . The case  $\theta = 0$ corresponds to our baseline in Section 3, where the planner acts exclusively on behalf of

 $<sup>^{31}</sup>$ Our model is closest to the baseline one-sector model in the working paper version of Midrigan and Xu (2010), who calibrate it to the South Korean development experience.

<sup>&</sup>lt;sup>32</sup>For comparison, South Korea's per capita GDP increased by a factor of about ten between 1970 and 2010. Of course, our model omits many of the real-world contributors to South Korea's growth, chief among them sustained productivity growth. Put differently, our calibration suggests that the TFP gains arising from financial deepening and improved capital allocation can alone account for over 20% of Korea's growth.

workers. The case  $\theta = 1$  corresponds to a utilitarian objective which weights all individuals equally, while in the intermediate case  $\theta = 1/2$  the planner weights each entrepreneur half as much as each worker.

#### 4.3 Growth and welfare with government interventions

We now study the growth and welfare consequences of various government policies. As in Section 3, we start with optimal labor taxes and then explore optimal credit subsidies to entrepreneurs. In addition, we contrast the results with various suboptimal policies, which may arise in practice due to political economy constraints.

We start by exploring the optimal labor tax schedules,  $\tau_{\ell}(t)$ , for different Pareto weights on entrepreneurs,  $\theta \in \{0, 1/2, 1\}$ , which we plot in the left panel of Figure 4. In all three cases, we recover our main result that optimal policy is *stage-dependent*, and the optimal labor tax in the beginning of the transition is lower than in the long run. In particular, a utilitarian planner who puts equal weight on all agents ( $\theta = 1$ ) would start the transition with a large labor supply subsidy of about 30%, and impose a labor tax in the long run equal to nearly 20%. Furthermore, the optimal policy subsidizes labor over an extended period of time, with the half-life of the policy equal to 13 years, and with 17 years before the subsidy is converted into a tax. This policy has a sizable effect on the GDP growth rates, increasing them from just over 5% on average under laissez-faire to 6% on average over the first 10 years of transition. This cumulates to a nearly 10% higher GDP in the 10th year of the policy relative to the laissez-faire, as we illustrate in the right panel of Figure 4, which plots the evolution of output under different policy regimes. Furthermore, we check that this pro-business policy results in a Pareto improvement and increases the welfare of both workers and entrepreneurs, as we discuss further below.

Next we consider the case with  $\theta < 1$ , so that the planner's objective in addition to efficiency also favors redistribution from the entrepreneurs towards the workers. As we can see from the left panel of Figure 4, reducing  $\theta$  uniformly shifts up the optimal tax schedule, consistent with our results in Section 3.2 (recall equation (32)). With  $\theta = 1/2$ , the planner still starts with a labor supply subsidy equal to 13%, which after 8 years turns into a labor tax that reaches nearly 21% in the long run. The right panel of Figure 4 shows that this greater preference for redistribution towards the workers reduces the growth rate of the economy relative to that achieved under the utilitarian policy with  $\theta = 1$ . Nonetheless, over the first 10 years the economy still grows faster than under laissez-faire, due to the



Figure 4: Optimal policy in the quantitative model

Note: Panel (a) plots the optimal labor tax schedules  $\tau_{\ell}(t)$  corresponding to different Pareto weights of entrepreneurs  $\theta \in \{0, 1/2, 1\}$ , as well as the labor tax imposed by a myopic labor union, as explained in the text. Panel (b) plots the evolution of GDP,  $Y(t) = \int y_t(a, z) d\mathcal{G}_t(a, z)$ , corresponding to the policies in panel (a), as well as the GDP dynamics under laissez-faire equilibrium, which is normalized to 1 in steady state.

optimal labor subsidy. In the long run, the optimal policy involves a labor tax, which acts to redistribute welfare from entrepreneurs to workers, and results in a lower long-run GDP than under laissez-faire.

When the planner cares exclusively about workers, putting a weight of  $\theta = 0$  on all entrepreneurs who together constitute a third of the population, the redistributive motive dominates the efficiency motive even early on in the transition. In this case, the planner starts with a positive, albeit tiny, labor tax, which increases over time to above 23%. This leads to uniformly lower growth rates than under the laissez-faire, yet workers gain at the expense of entrepreneurs. This, however, is not a general result when  $\theta = 0$ , and it depends on the initial condition for the wealth distribution of entrepreneurs. For example, if the economy starts out with even more undercapitalized entrepreneurs, with wealth levels scaled down 20-fold rather than 10-fold relative to the long run, the planner chooses to subsidize labor supply to entrepreneurs even when  $\theta = 0$ , echoing our theoretical results in Proposition 1.

To summarize, our quantitative analysis confirms that pro-business labor market policies in the early transition are optimal, even when the planner puts little weight on the well-being of entrepreneurs. To emphasize the importance of stage-dependent pro-business policies, we contrast the results with a particular form of *pro-labor policies*, namely a labor tax chosen by a *myopic labor union*. In particular, at each point in time t, the labor union maximizes the period utility of the workers,  $u(c(t), \ell(t))$ , without taking into account the equilibrium effects

	Welfare gains (%)		Annualized GDP growth $(\%)$		
	Overall	Workers	5 years	10 years	20 years
Optimal labor tax	0.503	0.751	9.59	5.51	3.01
Optimal flat labor tax, $\hat{\tau}_{\ell} = 0.100$	0.241	0.679	8.36	5.13	2.99
Flat labor subsidy, $-\underline{\tau}_{\ell} = 0.128$	-0.896	-1.442	10.16	6.05	3.46
Myopic union tax	-1.984	-0.378	5.38	3.65	2.28
Optimal credit subsidy	1.533	1.480	9.79	5.68	3.04

Table 1: Welfare and growth effects of policies

Note: Welfare gains in consumption-equivalent terms (i.e., % increase in consumption in every period).

that this policy has on wealth accumulation and endogenous productivity dynamics.<sup>33</sup> We derive the optimal union tax schedule in Appendix A5, and plot it in Figure 4 along with the resulting GDP dynamics under this tax. One important feature is that the union tax starts out high at 38% and decreases over time to 33% in the long run, in contrast with the optimal labor tax which starts low and increases over time. The reason is that the union tax does not factor in the dynamic efficiency consideration and simply depends on the elasticity of aggregate labor demand. The more constrained entrepreneurs are, the lower is the elasticity of labor demand, as they cannot adjust capital when financial constraints bind.<sup>34</sup> Thus, the same financial frictions that make pro-business policies optimal result in high labor taxes under a myopic labor union. As a consequence, this static union policy is detrimental to both GDP growth and welfare (see Table 1 discussed below). Even though this policy by design maximizes workers' current utility, it ends up being detrimental to workers, emphasizing the possible benevolent effects of the pro-business policies early on in the transition and the potential costs of the pro-labor policies in the financially-constrained economies.

We summarize the growth and welfare effects of various policies in Table 1. We use a standard consumption-equivalent welfare metric, that is the percentage change in consumption one would have to give individuals in the laissez-faire equilibrium each year to make them as well off as under the alternative policies (see Appendix A5 for details). For brevity, we focus on the intermediate case with the welfare weight on entrepreneurs  $\theta = 1/2$ . The

<sup>&</sup>lt;sup>33</sup>Such union policy can also proxy for other sources of labor market imperfections, which reduce equilibrium employment and increase labor costs to the firms, such as firing restrictions and severance payments, common in developing countries with continental-European labor market institutions (see e.g. Botero, Djankov, La Porta, Lopez-de Silanes, and Shleifer, 2004; Helpman and Itskhoki, 2010).

<sup>&</sup>lt;sup>34</sup>In the analytical model in Section 3, the optimal union tax is always equal to  $\gamma$ , as all active entrepreneurs operate at the borrowing constraint without ever growing out of it. Since the quantitative model features decreasing returns to scale, this is no longer the case, and the fraction of the constrained entrepreneurs deceases as the economy develops. In particular, under the laissez-faire, this fraction falls from 44% initially to 28% in the long run.

table reveals that the optimal labor tax in this case, which is pro-business in the short run and pro-worker in the long run, increases the combined welfare of workers and entrepreneurs by 0.5% in consumption-equivalent units. That is, the welfare increase from this policy is equivalent to the effect of a 0.5% increase in consumption in every period for every agent in the economy. Importantly, the optimal policy increases worker welfare by even more, namely by 0.75% in consumption equivalent. These numbers can be contrasted with those in the literature estimating the welfare cost of business cycles which are typically on the order of 0.01% (see e.g. Lucas, 2003). That is, the welfare gains from the optimal development policies are at least one order of magnitude larger than those from eliminating the business cycle. The table also shows that for developing countries the potential welfare and growth losses from myopic labor union policy are even larger.

To understand how important it is for policies to be stage-dependent, we also consider the welfare effects of various policies with time-invariant (flat) taxes. First, we consider an optimal flat tax. That is, we solve the same problem as above but under the restriction that  $\tau_{\ell}(t) = \hat{\tau}_{\ell}$  is constant for all t. The optimal flat tax is  $\hat{\tau}_{\ell} = 10\%$  and the resulting welfare gain is only 0.24%, i.e. less than half of the welfare gain under the optimal policy.<sup>35</sup> While this policy has only modest losses for workers relative to the best labor tax policy, the costs of this policy for GDP growth are considerably larger, resulting in 4% lower GDP after 10 years. Second, we consider the case in which the tax (subsidy) rate is set at  $\tau_{\ell} = -13\%$  and is then never adjusted, reflecting the possible power capture by organized lobbying groups. Such policy capture has large welfare costs for workers after short-run positive growth effects (*cf.* Buera, Moll, and Shin, 2013). These two results emphasize that the ability to subsidize labor supply to entrepreneurs early on in the transition is essential to ensure maximum welfare gains for both society at large and workers separately, but only provided that this policy is reversed when the economy becomes sufficiently developed.

Lastly, we study the optimal credit subsidy, which we parameterize analogously to the labor tax in (38). In the case  $\theta = 1/2$ , the optimal credit subsidy drops from an initial value of 100% to a long-run value of -70% with a half life of 7 years. The last row of Table 1 reports the resulting welfare effects, which are large and positive for both workers and entrepreneurs. The optimal credit subsidy mildly speeds up economic growth, yet doubles the welfare gains for workers and triples the welfare gains for the economy as a whole when compared with the optimal labor tax.<sup>36</sup> This echoes our Proposition 2, which emphasizes that in a constrained

<sup>&</sup>lt;sup>35</sup>Note that the optimal flat tax is different from the optimal steady-state tax, which does not take into account the welfare effects of transition, by analogy with the *golden rule* savings rate in capital accumulation. <sup>36</sup>This is, in part, the case because of the extreme optimal values of the credit subsidy in the short run

economy the planner would choose to use all available policy instruments to help the economy build entrepreneurial net worth in the short run and later use taxes to redistribute from entrepreneurs to workers to maximize their welfare gains.

Taken together, the results in this section again confirm our main message that probusiness policies are especially important for growth at earlier stages of development, and that such policies can be welfare-improving even from workers' perspective.

### 5 Optimal Policy in a Multi-Sector Economy

We now extend our analysis to a multi-sector environment. This allows us to study the optimal industrial policies and address a number of popular policy issues, such as promotion of comparative advantage sectors, optimal exchange rate policy, and infant industry protection. We summarize our main results here and provide the details of the environment and derivations in Appendix A6.

We assume households have general preferences  $u = u(c_0, c_1, \ldots, c_n)$  over n + 1 goods (sectors). Good i = 0 is an internationally-traded numeraire good with price normalized to  $p_0 = 1$ . Any of the remaining  $i \in \{1, ..., n\}$  goods can be either traded (T) or non-traded (N) internationally, and we denote their equilibrium (producer) prices with  $p_i$ . Traded good prices are taken as given in the international market  $(p_i = p_i^* \text{ for } i \in T)$ , while non-traded good prices are determined to clear the domestic market  $(c_i = y_i \text{ for } i \in N)$ . We further assume, for simplicity, that households supply L units of labor inelastically, and we study the allocation of aggregate labor supply across sectors,  $\sum_{i=0}^{n} \ell_i = L$ .

The main assumption that we make is that in each sector i, production expertise is entirely in the hands of specialized entrepreneurs, who hold aggregate sectoral wealth  $x_i$  and who are subject to financial frictions as described in Section 2. Lemma 2 generalizes in this case to the multi-sector environment, with (nominal) sectoral output given by

$$p_i y_i = p_i^{\zeta} \Theta_i x_i^{\gamma} \ell_i^{1-\gamma}, \quad \text{where} \quad \zeta \equiv 1 + \gamma(\eta - 1),$$

and sectoral wage rates given by

$$w_i = (1 - \alpha) \frac{p_i y_i}{\ell_i}, \qquad i \in \{0, 1, ..., n\}.$$

and the credit tax in the long run, which act as an effective way to redistribution between workers and entrepreneurs. If, however, the maximum sizes of the tax and the subsidy are capped, say at 50%, the quantitative welfare effects of such policy are much more in line with those of the optimal labor tax.
Sectoral productivity  $\Theta_i$  is defined as before, and may vary due to physical productivity  $A_i$  or financial constraints  $\lambda_i$ , which for example depend on the pledgeability of sectoral assets (see e.g. Rajan and Zingales, 1998; Manova, 2013).

We first study a planner that has access to sectoral labor income and consumption taxes,  $\{\tau_i^{\ell}, \tau_i^{c}\}_{i=0}^{n}$ , such that the after-tax (consumer) prices are  $\tilde{p}_i = (1 + \tau_i^{c})p_i$  and the after-tax wage rate is  $w = (1 - \tau_i^{\ell})w_i$ , equalized across sectors so that workers are indifferent about which sector to work in. We can also define an *overall sectoral wedge*,  $1 - \tau_i \equiv (1 - \tau_i^{\ell})/(1 + \tau_i^{c})$ , which summarizes the distortions that arise from both labor and consumption taxes.

Using the expressions above, we can solve for the sectoral labor allocation

$$\ell_i = \frac{\left((1-\tau_i^\ell)p_i^\zeta \Theta_i x_i^\gamma\right)^{1/\gamma}}{\sum_{j=0}^n \left((1-\tau_j^\ell)p_j^\zeta \Theta_j x_j^\gamma\right)^{1/\gamma}}L,\tag{40}$$

which we now study under various policy regimes. Note that labor taxes affect the sectoral labor allocation directly, while consumption taxes affect it indirectly, by changing the equilibrium producer prices  $p_i$ .

**Laissez-faire** In laissez-faire equilibrium, with no taxes  $\tau_i^{\ell} = \tau_i^c \equiv 0$ , the equilibrium sectoral labor shares are proportional to the labor productivity shifters  $p_i^{\zeta} \Theta_i x_i^{\gamma}$ , which depend in part on the accumulated financial wealth of the sectoral entrepreneurs. In the long run, financial wealth is endogenously accumulated and reflects the fundamental sectoral productivity  $p_i^{\zeta} \Theta_i$ . Therefore, the long-run laissez-faire labor allocation does not depend on the initial wealth distribution across sectors  $\{x_i(0)\}$ , which however is important in shaping the allocations along the transition path.

**Optimal policy interventions** Our theoretical results in Appendix A6 emphasize two main principles of the optimal sectoral policies:

1. zero consumption taxes in the tradable sectors, and labor subsidies to relax the sectoral financial constraints (as in a one-sector economy, cf. (30)):

$$\tau_i^c = 0 \quad \text{and} \quad \tau_i^\ell = \gamma(1 - \nu_i) \quad \text{for} \quad i \in T,$$
(41)

2. zero overall sectoral wedges (as defined above) in the non-tradable sectors:

$$\tau_i = 0 \quad \text{with} \quad \tau_i^c = -\tau_i^\ell = \frac{1}{\eta - 1}(\nu_i - 1) \quad \text{for} \quad i \in N,$$
(42)

where in both cases  $\nu_i$  is the shadow value of entrepreneurial wealth in sector *i*.



Figure 5: Planner's allocation in an economy with two tradable sectors

Note: The sectors are symmetric in all but their latent comparative advantage, with  $p_0^{*\zeta}\Theta_0 > p_1^{*\zeta}\Theta_1$ . Panel (a) plots the labor supply subsidy to the comparative advantage sector 0. Panel (b) plots the evolution of the sectoral entrepreneurial wealth under laissez-faire (dashed lines) and optimal sectoral labor taxes (solid lines).

In a small open economy, the planner chooses not to manipulate consumption prices of tradable goods, as this cannot increase the profitability of the domestic producers due to perfectly elastic foreign supply. The planner instead subsidizes labor reallocation towards the tradable sectors with high shadow value of financial wealth  $\nu_i$ , i.e. the sectors that are undercapitalized relative to their fundamental productivity. In contrast, for non-tradable goods, the planner chooses to manipulate equilibrium prices  $p_i$  using consumption taxes, offsetting the resulting sectoral wedges with labor subsidies. This is indeed the least distortive way to increase the profitability of the non-tradable sectors with high shadow values of entrepreneurial wealth.

We consider next three special cases, which illustrate these general principles:

Comparative advantage and industrial policies The most immediate application of our results is to economy with tradable sectors only. In this case, the planner simply tilts the allocation of labor across sectors according to the shadow values of entrepreneurial wealth  $\nu_i$  by means of sectoral labor taxes (see (40) and (41)). This relaxes, over time, the financial constraints that bind the most in the economy.<sup>37</sup> We further show that, for a given level of entrepreneurial wealth  $x_i$ , its shadow value  $\nu_i$  increases with the *latent*, or long-run, *compara*-

<sup>&</sup>lt;sup>37</sup>This policy can only be second-best, as it distorts the equalization of marginal products of labor across sectors. In Appendix A6 we generalize this analysis, along the lines of Proposition 2, to allow for additional sectoral policy instruments, including production, credit and export subsidies.

tive advantage of the sector, as captured by the revenue productivity  $p_i^{*\zeta}\Theta_i$ . A sector's actual, or short-run, comparative advantage  $p_i^{*\zeta}\Theta_i x_i^{\gamma}$  differs from its latent comparative advantage, and depends on accumulated sectoral wealth. In the short run, the country may specialize against its latent comparative advantage, if entrepreneurs in those sectors are poorly capitalized (see Wynne, 2005). Therefore, the planner tilts sectoral labor allocation towards the long-run latent comparative advantage sectors, and hence speeds up the transition in this open economy, as illustrated in Figure 5. This implication of our analysis is consistent with some popular policy prescriptions; however, identifying the latent comparative advantage sectors may be a challenging task in practice (see e.g. Stiglitz and Yusuf, 2001, as well as two empirical approaches to this challenge in Hidalgo, Klinger, Barabási, and Hausmann, 2007 and in Lin, 2012).

Real exchange rate and competitiveness Consider next a two-sector model with a tradable sector i = 0 and a non-tradable sector i = 1, which allows us to study the real exchange rate implications of the optimal policy. In this economy, the consumption-based real exchange rate is defined by the effective consumer price of non-tradables,  $(1 + \tau_1^c)p_1$ .<sup>38</sup> Specializing the general optimal policy characterization in (41)–(42) to this case, we see that the planner subsidizes the labor supply to the tradable sector i = 0 whenever  $\nu_0 > 1$ , independently of the tightness of the financial constraints in the non-tradable sector. Hence labor is diverted away from non-tradables to tradables and, since production features decreasing returns to labor, equilibrium labor costs in the tradable sector  $w_0 = (1-\alpha_0)y_0/\ell_0$  are compressed, increasing the international competitiveness of the economy. In contrast, this increases relative labor costs in the non-tradable sector, and hence leads to an *appreciated* consumer-price real exchange rate due to more expensive non-tradable goods.<sup>39</sup>

The situation is different when the planner does not have access to any sectoral taxes and has to resort to intertemporal distortions by means of a savings subsidy, or a policy of capital controls and reserve accumulation more commonly used in practice (see Jeanne, 2013, for the equivalence result of these policies). By taxing consumption today in favor of future

<sup>&</sup>lt;sup>38</sup>The CPI-based real exchange rate is given by  $P/P^*$ , where P and  $P^*$  are the price indexes of the home country and the rest of the world which are functions of the consumer prices of tradable and non-tradable goods. Since we analyze a small open economy,  $P^*$  is fixed from the point of view of the home country, and we normalize  $p_0 = 1$  and  $\tau_0^c = 0$ . Therefore, the real exchange rate appreciates whenever the consumer price of non-tradables  $(1 + \tau_1^c)p_1$  increases.

<sup>&</sup>lt;sup>39</sup>Furthermore, if  $\nu_1 > 1$ , the planner subsidizes the non-tradable producers by increasing the equilibrium price of non-tradables using a consumption tax, further appreciating the real exchange rate. In Appendix A6, we generalize this result to the case when the planner cannot directly tax sectoral labor, as distinguishing between tradable and non-tradable labor may be difficult, and can only tax sectoral consumption.

periods, the planner shifts resources away from the non-tradable sector and towards the tradable sector, which is desirable when  $\nu_0$  is sufficiently large. As a result, wages and prices of non-tradables as well as consumption of both goods decrease, while the tradable sector expands production and exports facing unchanged international prices.<sup>40</sup> In this case, greater competitiveness of the country in the tradable sector is indeed associated with cheaper non-tradables and a depreciated real exchange rate. This policy, however, induces an unnecessary intertemporal distortion, and hence is at most third-best and is strictly dominated whenever static sectoral taxes are available. To summarize, while the goal of the planner may be to compress wages and shift labor towards the tradable sector, the implications for the real exchange rate are sensitive to the set of the available policy instruments, making it an inconvenient target for policymakers (*cf.* Rodrik, 2008).

Cohorts of entrepreneurs and infant industry protection Lastly, we consider a generalization of the baseline model with overlapping generations of cohorts of entrepreneurs. We show in Appendix A6 that the optimal multi-sector policy rule (41) still applies in this case. Specifically, instead of sectors, *i* now refers to the date of birth of the cohort of entrepreneurs, and  $p_i \equiv 1$  for all *i* since we assume that all entrepreneurs produce the same international numeraire good. What makes this setup interesting is if the new cohorts of entrepreneurs have higher levels of productivity, e.g. come in with new ideas, captured with an increasing profile of  $\Theta_i$  with *i*. At the same time, the young entrepreneurs enter undercapitalized relative to the average existing entrepreneurs in the economy, who have been accumulating financial wealth from their past profits. By analogy with the multi-sector economy, the planner chooses to subsidize the employment of the younger cohorts of entrepreneurs, which is reminiscent of infant industry protection policies, albeit for different reasons than typically put forward (*cf.* Corden, 1997, chapter 8).

## 6 Conclusion

The presence of financial frictions opens the door for welfare-improving government interventions in product and factor markets. We develop a framework to study the Ramsey-optimal interventions which improve welfare and accelerate economic development in financially underdeveloped economies. The main insight of our analysis is that dynamic stage-dependent pro-business policies can generically improve welfare, including that of workers. For example,

<sup>&</sup>lt;sup>40</sup>Interestingly, this narrative is consistent with the analysis in Song, Storesletten, and Zilibotti (2014) who argue that in China a combination of capital controls and other policies compressed wages and increased the wealth of entrepreneurs, thereby relaxing their borrowing constraints.

financial frictions justify a policy intervention that increases labor supply and reduces wages in the early stages of transition so as to speed up entrepreneurial wealth accumulation and relax future financial constraints, which in turn leads to higher labor productivity and wages. However, the optimal policy reverses sign along the transition and becomes pro-worker in the long run. More generally, the optimal policy mix also includes credit and production subsidies, all combined together in a pro-business fashion in the early transition, and then reversed in favor of more redistributive goals later on.

To facilitate the analysis, we develop a particularly tractable version of the workhorse macro-development growth model with heterogeneous entrepreneurs facing financial constraints. This tractability allows for a sharp analytical characterization of the optimal policies along the transition path of the economy. It also allows us to consider a number of extensions, for example to an environment with overlapping generations of finitely-lived workers and entrepreneurs facing similar borrowing constraints. In addition, we can study optimal policies in an environment with multiple tradable and non-tradable sectors, addressing the desirability of various popular industrial and exchange rate policies. Our baseline model relies on a number of strong assumptions, which we relax in our quantitative analysis, thereby confirming the robustness of our findings and the quantitative relevance of the policies we focus on for growth and welfare. Our normative analysis provides an efficiency rationale, but also identifies caveats, for many of the development policies actively pursued by dynamic emerging economies.

## References

- ACEMOGLU, D., P. AGHION, AND F. ZILIBOTTI (2006): "Distance to Frontier, Selection, and Economic Growth," Journal of the European Economic Association, 4(1), 37–74.
- ACHDOU, Y., J. HAN, J.-M. LASRY, P.-L. LIONS, AND B. MOLL (2017): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," NBER Working Paper 23732.
- AGHION, P., AND P. BOLTON (1997): "A Theory of Trickle-Down Growth and Development," *The Review of Economic Studies*, 64, 151–172.
- AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," The Quarterly Journal of Economics, 109(3), 659–84.
- (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103(6), 1158–75.
- AMARAL, P. S., AND E. QUINTIN (2010): "Limited Enforcement, Financial Intermediation, And Economic Development: A Quantitative Assessment," *International Economic Review*, 51(3), 785–811.
- ANGELETOS, G.-M., F. COLLARD, H. DELLAS, AND B. DIBA (2013): "Optimal Public Debt Management and Liquidity Provision," NBER Working Papers 18800.
- ASKER, J., A. COLLARD-WEXLER, AND J. DE LOECKER (2014): "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013–1063.
- ATKESON, A., AND P. J. KEHOE (2007): "Modeling the Transition to a New Economy: Lessons from Two Technological Revolutions," *American Economic Review*, 97(1), 64–88.
- BACCHETTA, P., K. BENHIMA, AND Y. KALANTZIS (2014): "Optimal Exchange Rate Policy in a Growing Semi-Open Economy," *IMF Economic Review*, 62(1), 48–76.
- BANERJEE, A. V., AND E. DUFLO (2005): "Growth Theory through the Lens of Development Economics," in Handbook of Economic Growth, ed. by P. Aghion, and S. Durlauf, vol. 1 of Handbook of Economic Growth, chap. 7, pp. 473–552. Elsevier.
- BANERJEE, A. V., AND A. F. NEWMAN (1993): "Occupational Choice and the Process of Development," The Journal of Political Economy, 101(2), 274–298.
  - (2003): "Inequality, Growth and Trade Policy," Working paper, MIT.
- BARRO, R. J. (1979): "On the Determination of the Public Debt," *Journal of Political Economy*, 87(5), 940–71.
- BECK, T., A. DEMIRGUC-KUNT, AND R. LEVINE (2000): "A New Database on the Structure and Development of the Financial Sector," World Bank Economic Review, 14(3), 597–605.
- BENIGNO, G., AND L. FORNARO (2012): "Reserve Accumulation, Growth and Financial Crises," CEPR Discussion Papers 9224.
- BIANCHI, J. (2011): "Overborrowing and Systemic Externalities in the Business Cycle," American Economic Review, 101(7), 3400–3426.
- BLANCHARD, O. J. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93(2), 223–47.
- BLUNDELL, R., L. PISTAFERRI, AND I. SAPORTA-EKSTEN (2016): "Consumption Inequality and Family Labor Supply," American Economic Review, 106(2), 387–435.
- BOTERO, J. C., S. DJANKOV, R. LA PORTA, F. LOPEZ-DE SILANES, AND A. SHLEIFER (2004): "The Regulation of Labor," *The Quarterly Journal of Economics*, 119(4), 1339–1382.
- BRUNNERMEIER, M. K., T. M. EISENBACH, AND Y. SANNIKOV (2013): "Macroeconomics with Financial Frictions: A Survey," in Advances in Economics and Econometrics, ed. by D. Acemoglu, M. Arellano, and E. Dekel, vol. II of Tenth World Congress of the Econometric Society, pp. 3–94. Cambridge University Press, Cambridge, UK.
- BUERA, F., AND J.-P. NICOLINI (2017): "Liquidity Traps and Monetary Policy: Managing a Credit Crunch," Federal Reserve Bank of Minneapolis, Staff Report 541.
- BUERA, F. J., J. P. KABOSKI, AND Y. SHIN (2011): "Finance and Development: A Tale of Two Sectors," American Economic Review, 101(5), 1964–2002.

(2012): "The Macroeconomics of Microfinance," NBER Working Papers No. 17905.

- BUERA, F. J., B. MOLL, AND Y. SHIN (2013): "Well-Intended Policies," *Review of Economic Dynamics*, 16(1), 216–230.
- BUERA, F. J., AND Y. SHIN (2013): "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 121(2), 221 272.
- CABALLERO, R. J., AND G. LORENZONI (2014): "Persistent Appreciations and Overshooting: A Normative Analysis," *IMF Economic Review*, 62(1), 1–47.
- CAGETTI, M., AND M. DE NARDI (2006): "Entrepreneurship, Frictions and Wealth," Journal of Political Economy, 114(5), 835–870.
- CALVO, G. A., AND M. OBSTFELD (1988): "Optimal Time-Consistent Fiscal Policy with Finite Lifetimes," *Econometrica*, 56(2), 411–32.
- CASELLI, F., AND N. GENNAIOLI (2013): "Dynastic Management," Economic Inquiry, 51(1), 971–996.
- CHETTY, R., A. GUREN, D. MANOLI, AND A. WEBER (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101(3), 471–475.
- COLE, H. L., AND L. E. OHANIAN (2004): "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis," *Journal of Political Economy*, 112(4), 779–816.
- CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): "Taxing Capital? Not a Bad Idea after All!," American Economic Review, 99(1), 25–48.
- CORDEN, W. M. (1997): Trade Policy and Economic Welfare. Clarendon Press, 2nd edn.
- COSTINOT, A., G. LORENZONI, AND I. WERNING (2014): "A Theory of Capital Controls as Dynamic Terms-of-Trade Manipulation," *Journal of Political Economy*, 122(1), 77–128.
- DÁVILA, J., J. HONG, P. KRUSELL, AND J.-V. RÍOS-RULL (2012): "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks," *Econometrica*, 80(6), 2431–2467.
- DE LOECKER, J., P. K. GOLDBERG, A. K. KHANDELWAL, AND N. PAVCNIK (2016): "Prices, Markups, and Trade Reform," *Econometrica*, 84, 445–510.
- DI TELLA, S., AND Y. SANNIKOV (2016): "Optimal Asset Management Contracts with Hidden Savings," https://web.stanford.edu/~sditella/research.html.
- DIAZ-ALEJANDRO, C. (1985): "Good-bye Financial Repression, Hello Financial Crash," Journal of Development Economics, 19(1), 1–24.
- EROSA, A., AND A. HIDALGO-CABRILLANA (2008): "On Finance As A Theory Of TFP, Cross-Industry Productivity Differences, And Economic Rents," *International Economic Review*, 49(2), 437–473.
- FEINSTEIN, C. H. (1998): "Pessimism perpetuated: real wages and the standard of living in Britain during and after the industrial revolution," *The Journal of Economic History*, 58(03), 625–658.
- GALOR, O., AND J. ZEIRA (1993): "Income Distribution and Macroeconomics," *The Review of Economic Studies*, 60(1), 35–52.
- GREENWALD, B. C., AND J. E. STIGLITZ (1986): "Externalities in Economies with Imperfect Information and Incomplete Markets," *The Quarterly Journal of Economics*, 101(2), 229–64.
- GROWTH COMMISSION (2008): The Growth Report: Strategies for Sustained Growth and Inclusive Development. Washington, DC: World Bank.
- HARRISON, A., AND A. RODRÍGUEZ-CLARE (2010): Trade, Foreign Investment, and Industrial Policy for Developing Countries pp. 4039–4214, Handbook of Development Economics. Elsevier.
- HELPMAN, E., AND O. ITSKHOKI (2010): "Labour Market Rigidities, Trade and Unemployment," *Review of Economic Studies*, 77, 1100–1137.
- HIDALGO, C. A., B. KLINGER, A.-L. BARABÁSI, AND R. HAUSMANN (2007): "The Product Space Conditions the Development of Nations," *Science*, 317(5837), 482–487.
- HSIEH, C.-T., AND P. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4), 1403–1448.
- JEANNE, O. (2013): "Capital Account Policies and the Real Exchange Rate," NBER International Seminar on Macroeconomics, 9(1), 7–42.

- JEANNE, O., AND A. KORINEK (2010): "Managing Credit Booms and Busts: A Pigouvian Taxation Approach," NBER Working Papers No. 16377.
- JEONG, H., AND R. TOWNSEND (2007): "Sources of TFP Growth: Occupational Choice and Financial Deepening," *Economic Theory*, 32, 179–221.
- JUDD, K. L. (1985): "Redistributive taxation in a simple perfect foresight model," Journal of Public Economics, 28(1), 59–83.
- KEHOE, T. J., AND D. K. LEVINE (2001): "Liquidity Constrained Markets versus Debt Constrained Markets," *Econometrica*, 69(3), 575–598.
- KIM, D.-I., AND R. H. TOPEL (1995): "Labor Markets and Economic Growth: Lessons from Korea's Industrialization, 1970-1990," in *Differences and Changes in Wage Structures*, pp. 227–264. National Bureau of Economic Research, Cambridge.
- KIM, K., AND D. M. LEIPZIGER (1997): "Korea: A Case of Government-Led Development," in Lessons from East Asia, ed. by D. M. Leipziger, pp. 155–212. University of Michigan Press, Ann Arbor.
- KIYOTAKI, N., AND J. MOORE (1997): "Credit Cycles," Journal of Political Economy, 105(2), 211-48.
- KORINEK, A., AND L. SERVÉN (2016): "Undervaluation through foreign reserve accumulation: Static losses, dynamic gains,," *Journal of International Money and Finance*, 64, 104–136.
- KRUGMAN, P. R. (1987): "The narrow moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: Notes on trade in the presence of dynamic scale economies," *Journal of Development Economics*, 27(1-2), 41–55.
- LEIPZIGER, D. M. (1997): Lessons from East Asia. University of Michigan Press, Ann Arbor, MI.
- LEVINE, R. (2005): "Finance and Growth: Theory and Evidence," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf. Elsevier B.V.
- LIN, J. Y. (2012): The Quest for Prosperity: How Developing Economies Can Take Off. Princeton, NJ: Princeton University Press.
- LORENZONI, G. (2008): "Inefficient Credit Booms," Review of Economic Studies, 75(3), 809-833.
- LUCAS, JR., R. E. (2003): "Macroeconomic Priorities," American Economic Review, 93(1), 1–14.
- LUCAS, JR., R. E., AND N. L. STOKEY (1983): "Optimal fiscal and monetary policy in an economy without capital," *Journal of Monetary Economics*, 12(1), 55–93.
- MANOVA, K. (2013): "Credit Constraints, Heterogeneous Firms, and International Trade," The Review of Economic Studies, 80(2), 711–744.
- MATSUYAMA, K. (1992): "Agricultural productivity, comparative advantage, and economic growth," *Journal* of Economic Theory, 58(2), 317–334.
- (2008): "Aggregate Implications of Credit Market Imperfections," in *NBER Macroeconomics Annual 2007*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, vol. 22, pp. 1–60. Chicago, II: University of Chicago Press.
- MCKINNON, R. I. (1981): "Financial Repression and the Liberalization Problem within Less Developed Countries," in *The World Economic Order: Past and Prospects*, ed. by S. Grassman, and E. Lundberg, pp. 365–386. St. Martin's Press, New York.
- MICHELACCI, C., AND V. QUADRINI (2009): "Financial Markets and Wages," *Review of Economic Studies*, 76(2), 795–827.
- MIDRIGAN, V., AND D. Y. XU (2010): "Finance and Misallocation: Evidence from Plant-level Data," Discussion paper, NBER Working Papers No. 15647.
- (2014): "Finance and Misallocation: Evidence from Plant-Level Data," American Economic Review, 104(2), 422–458.
- MOLL, B. (2014): "Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation?," American Economic Review, 104(10), 3186–3221.
- MURPHY, K. M., A. SHLEIFER, AND R. W. VISHNY (1989): "Industrialization and the Big Push," *Journal* of *Political Economy*, 97(5), 1003–26.
- NUÑO, G., AND B. MOLL (2018): "Social Optima in Economies with Heterogeneous Agents," Review of Economic Dynamics, 28, 150–180.

- RAJAN, R. G., AND L. ZINGALES (1998): "Financial Dependence and Growth," The American Economic Review, 88(3), 559–586.
- RESTUCCIA, D., AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11(4), 707–720.
- RODRIK, D. (2008): "The Real Exchange Rate and Economic Growth," Brookings Papers on Economic Activity, 2008(2), 365–412.
- SCHUMAN, M. (2010): The Miracle: The Epic Story of Asia's Quest for Wealth. Harper Collins.
- SHIN, Y. (2006): "Ramsey Meets Bewley: Optimal Government Financing with Incomplete Markets," https://pages.wustl.edu/yshin/research.
- SONG, Z., K. STORESLETTEN, AND F. ZILIBOTTI (2011): "Growing Like China," American Economic Review, 101(1), 196–233.
  - (2014): "Growing (with Capital Controls) like China," *IMF Economic Review*, 62(3), 327–370.
- STIGLITZ, J. E., AND S. YUSUF (2001): Rethinking the East Asian Miracle. World Bank.
- STOKEY, N. L. (2009): The Economics of Inaction. Princeton, NJ: Princeton University Press.
- STRAUB, L., AND I. WERNING (2014): "Positive Long Run Capital Taxation: Chamley-Judd Revisited," NBER Working Papers No. 20441.
- TINBERGEN, J. (1984): "Optimal Development Policies: Lessons from Experience," *The World Economy*, 7(1), 112–117.
- TOWNSEND, R. M. (2010): "Financial Structure and Economic Welfare: Applied General Equilibrium Development Economics," Annual Review of Economics, 2, 507–546.
- VENTURA, J., AND H.-J. VOTH (2015): "Debt into Growth: How Sovereign Debt Accelerated the First Industrial Revolution," NBER Working Paper No. 21280.
- VOTH, H.-J. (2001): "The longest years: new estimates of labor input in England, 1760–1830," *The Journal of Economic History*, 61(04), 1065–1082.
- WEITZMAN, M. L. (1974): "Prices vs. Quantities," Review of Economic Studies, 41(4), 477–91.
- WERNING, I. (2007): "Optimal Fiscal Policy with Redistribution," The Quarterly Journal of Economics, 122(3), 925–967.
- WYNNE, J. (2005): "Wealth as a Determinant of Comparative Advantage," American Economic Review, 95(1), 226–254.
- YAARI, M. E. (1965): "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," Review of Economic Studies, 32(2), 137–150.
- YOUNG, A. (1991): "Learning by Doing and the Dynamic Effects of International Trade," *The Quarterly Journal of Economics*, 106(2), 369–405.

## Online Appendix A "Optimal Development Policies with Financial Frictions"

Forthcoming in <u>Econometrica</u>

Oleg Itskhoki and Benjamin Moll Princeton University

July 24, 2018

# A Derivations and Proofs

## Contents

Δ1	Derivations and Proofs for Section 2		
AI			2
	A1.1		2
	A1.2	Proofs of Lemmas 1 and 2	2
	A1.3	Entrepreneurs: value and policy functions	3
A2	An Economy with Transfers		5
	A2.1	Desirability of transfers from workers to entrepreneurs	5
	A2.2	Optimal policy with transfers to entrepreneurs	6
	A2.3	Infeasibility of transfers	10
A3	Derivations and Proofs for Section 3		11
	A3.1	Optimality conditions for the planner's problem	11
	A3.2	Proof of Proposition 1	13
	A3.3	Additional tax instruments	15
	A3.4	Finite lives and financially constrained households	19
A4	Exten	sions	22
	A4.1	Persistent productivity types	22
	A4.2	Pareto weight on entrepreneurs	23
	A4.3	Closed economy	24
	A4.4	Optimal intertemporal wedge	26
A5	Apper	ndix for Quantitative Model in Section 4	27
A6	Analysis of the Multi-sector Model in Section 5		33
	A6.1	Setup of the multi-sector economy	33
	A6.2	Optimal policy	35
	A6.3	Comparative advantage and industrial policies	37
	A6.4	Non-tradables, the real exchange rate and competitiveness	39
	A6.5	Optimality conditions in the multi-sector economy	41
	A6.6	Overlapping production cohorts	45

## A1 Derivations and Proofs for Section 2

### A1.1 Frisch labor supply elasticity

For any utility function  $u(c, \ell)$  defined over consumption c and labor  $\ell$ , consider the system of equations

$$u_c(c,\ell) = \mu,\tag{A1}$$

$$u_{\ell}(c,\ell) = -\mu w. \tag{A2}$$

These two equations define  $\ell$  and c as a function of the marginal utility  $\mu$  and the wage rate w. The solution for  $\ell$  is called the *Frisch labor supply* function and we denote it by  $\ell = \ell^F(\mu, w)$ . We assumed that the utility function features a positive and finite Frisch labor supply elasticity for all  $(\mu, w)$ :

$$\varepsilon(\mu, w) \equiv \frac{\partial \log \ell^F(\mu, w)}{\partial \log w} = \frac{1}{\frac{u_{\ell\ell}\ell}{u_{\ell}} - \frac{(u_{c\ell})^2\ell}{u_{cr}u_{\ell}}} \in (0, \infty),$$
(A3)

where the second equality comes from a full differential of (A1)–(A2) under constant  $\mu$ , which we simplify using  $w = -u_{\ell}/u_c$  implied by the ratio of (A1) and (A2). Therefore, the condition we impose on the utility function is:

$$\frac{u_{\ell\ell}\ell}{u_{\ell}} > \frac{(u_{c\ell})^2\ell}{u_{cc}u_{\ell}} \qquad \Leftrightarrow \qquad u_{\ell\ell}u_{cc} > (u_{c\ell})^2 \tag{A4}$$

for all possible pairs  $(c, \ell)$ . Due to convexity of  $u(\cdot)$ , this in particular implies  $u_{\ell\ell} < 0$ .

## A1.2 Proofs of Lemmas 1 and 2

**Proof of Lemma 1** Equation (9) is the first-order condition of profit maximization  $\pi(a, z)$  in (6) with respect to n, which substituted into the profit function results in:

$$\pi(a,z) = \max_{0 \le k \le \lambda a} \left\{ \left( \alpha \left[ (1-\alpha)/w \right]^{(1-\alpha)/\alpha} A^{1/\alpha} z - r^* \right) k \right\}.$$

Equations (8) and (11) characterize the solution to this problem of maximizing a linear function of k subject to inequality constraints  $0 \le k \le \lambda a$ . Finally, we substitute (11) into the expression for profits to obtain (10). The assumption that the least productive entrepreneur is inactive along the full transition path and for any initial conditions can be ensured by choosing a sufficient amount of productivity heterogeneity ( $\eta$  small enough).<sup>1</sup>

**Aggregation** We next provide derivations for equations (13)–(15) in the text:

$$\kappa = \int k_t(a, z) \mathrm{d}\mathcal{G}_t(a, z) = \int_{z \ge \underline{z}} \left[ \int \lambda a \, \mathrm{d}G_{a, t}(a) \right] \mathrm{d}G_z(z) = \lambda x \left[ 1 - G_z(\underline{z}) \right] = \lambda x \underline{z}^{-\eta}$$

<sup>&</sup>lt;sup>1</sup>However, in the limit without heterogeneity  $(\eta \to \infty)$ , this assumption is necessarily violated, yet the analysis of the case when all entrepreneurs produce  $(\underline{z} = 1)$  yields similar qualitative results at the cost of some additional notation.

and

$$\ell = \int n_t(a, z) \mathrm{d}\mathcal{G}_t(a, z) = \left[ (1 - \alpha)A/w \right]^{1/\alpha} \int_{z \ge \underline{z}} z \left[ \int \lambda a \, \mathrm{d}G_{a, t}(a) \right] \mathrm{d}G_z(z)$$
$$= \left[ (1 - \alpha)A/w \right]^{1/\alpha} \lambda x \left[ 1 - G_z(\underline{z}) \right] \mathbb{E} \{ z | z \ge \underline{z} \} = \left[ (1 - \alpha)A/w \right]^{1/\alpha} \lambda x \frac{\eta}{\eta - 1} \underline{z}^{1 - \eta}$$

where we substitute the policy functions (8)–(9) into the definitions of  $\kappa$  and  $\ell$ , and then integrate making use of the independence of the *a* and *z* distributions, the definition of aggregate wealth *x*, and the Pareto distribution assumption for *z*. Similarly, we calculate:

$$y = \int A(zk_t(a,z))^{\alpha} n_t(a,z)^{1-\alpha} \mathrm{d}\mathcal{G}_t(a,z) = A[(1-\alpha)A/w]^{\frac{1-\alpha}{\alpha}} \int_{z \ge \underline{z}} z \left[ \int \lambda a \, \mathrm{d}G_{a,t}(a) \right] \mathrm{d}G_z(z)$$
$$= A\kappa^{\alpha} \ell^{1-\alpha} \left(\frac{\eta}{\eta-1}\underline{z}\right)^{\alpha},$$

where we isolate out the  $\kappa$  and  $\ell$  terms on the right-hand side and the last term in brackets emerges as a residual.

**Proof of Lemma 2** Combine cutoff condition (11) and labor demand (14), and solve out the wage rate w to obtain the expression for cutoff  $\underline{z}$  in (17). Substitute the resulting expression (17) and capital demand (13) into the aggregate production function to obtain expression (16) for aggregate output y as a function of  $\ell$  and x. The remaining equations are a result of direct manipulation of (13)–(15) and (17), after noting that aggregate profits are an integral of individual profits in (10) and equal to:

$$\Pi = \int \left(\frac{z}{\underline{z}} - 1\right) r^* k_t(a, z) \mathrm{d}\mathcal{G}_t(a, z) = r^* \int_{z \ge \underline{z}} \left(\frac{z}{\underline{z}} - 1\right) \left[\int \lambda a \, \mathrm{d}G_{a, t}(a)\right] \mathrm{d}G_z(z) = \frac{r^* \kappa}{\eta - 1}.$$

#### A1.3 Entrepreneurs: value and policy functions

**Lemma A4** Consider an entrepreneur with logarithmic utility, discount factor  $\delta$ , and budget constraint  $\dot{a} = R_t(z)a - c_e$  for some  $R_t(z)$ , where z is iid over time. Then her consumption policy function is  $c_e = \delta a$  and her expected value starting from initial assets  $a_0$  is

$$V_0(a_0) = -\frac{1}{\delta}(1 - \log \delta) + \frac{1}{\delta}\log a_0 + \frac{1}{\delta}\int_0^\infty e^{-\delta t} \mathbb{E}_z R_t(z) \mathrm{d}t.$$
 (A5)

**Proof**: This derivation follows the proof of Lemma 2 in Moll (2014). Denote by  $v_t(a, z)$  the value to an entrepreneur with assets a and productivity z at time t, which can be expressed recursively as (see Chapter 2 in Stokey, 2009):

$$\delta v_t(a,z) = \max_{c_e} \left\{ \log c_e + \frac{1}{\mathrm{d}t} \mathbb{E} \{ \mathrm{d}v_t(a,z) \}, \quad \text{s.t. } \mathrm{d}a = [R_t(z)a - c_e] \mathrm{d}t \right\}.$$

The value function depends on calendar time t because prices and taxes vary over time. In the absence of aggregate shocks, from the point of view of entrepreneurs, calendar time is a "sufficient statistic" for the evolution of the distribution  $\mathcal{G}_t(a, z)$ .

The proof proceeds with a guess and verify strategy. Guess that the value function takes the form  $v_t(a, z) = B\tilde{v}_t(z) + B\log a$ . Using this guess we have that  $\mathbb{E}\{dv_t(a, z)\} = Bda/a + B\mathbb{E}\{d\tilde{v}_t(z)\}$ . Rewrite the value function:

$$\delta B\tilde{v}_t(z) + \delta B\log a = \max_{c_e} \left\{ \log c_e + \frac{B}{a} \left[ R_t(z)a - c_e \right] + B \frac{1}{dt} \mathbb{E} \{ \mathrm{d}\tilde{v}_t(z) \} \right\}.$$

Take first-order condition to obtain  $c_e = a/B$ . Substituting back in,

$$\delta B\tilde{v}_t(z) + \delta B\log a = \log a - \log B + BR_t(z) - 1 + B\frac{1}{dt}\mathbb{E}\{\mathrm{d}\tilde{v}_t(z)\}.$$

Collecting the terms involving log a, we see that  $B = 1/\delta$  so that  $c_e = \delta a$  and  $\dot{a} = [R_t(z) - \delta]a$ , as claimed in (12) in the text.

Finally, the value function is

$$v_t(a,z) = \frac{1}{\delta} \big( \tilde{v}_t(z) + \log a \big), \tag{A6}$$

confirming the initial conjecture, where  $\tilde{v}_t(z)$  satisfies

$$\delta \tilde{v}_t(z) = \delta(\log \delta - 1) + R_t(z) + \frac{1}{dt} \mathbb{E}\{\mathrm{d}\tilde{v}_t(z)\}.$$
(A7)

Next we calculate the expected value:

$$V_0(a_0) = \int v_0(a_0, z) g_z(z) dz = \frac{1}{\delta} \left( \tilde{V}_0 + \log a_0 \right),$$

where  $g_z(\cdot)$  is the pdf of z and  $\tilde{V}_0 \equiv \int \tilde{v}_0(z)g_z(z)dz$ . Integrating (A7):

$$\delta \tilde{V}_t = \delta(\log \delta - 1) + \int R_t(z)g_z(z)dz + \dot{\tilde{V}}_t, \tag{A8}$$

where we have used that (under regularity conditions so that we can exchange the order of integration)

$$\int \frac{1}{dt} \mathbb{E} \{ \mathrm{d}\tilde{v}_t(z) \} g_z(z) \mathrm{d}z = \frac{1}{dt} \mathbb{E} \left\{ \mathrm{d} \int \tilde{v}_t(z) g_z(z) \mathrm{d}z \right\} = \frac{1}{dt} \mathbb{E} \{ \mathrm{d}\tilde{V}_t \} = \dot{\tilde{V}}_t.$$

Integrating (A8) forward in time:

$$\tilde{V}_0 = \log \delta - 1 + \int_0^\infty e^{-\delta t} \left[ \int R_t(z) g_z(z) dz \right] dt,$$

and hence

$$V_0(a_0) = -\frac{1}{\delta}(1 - \log \delta) + \frac{1}{\delta}\log a_0 + \frac{1}{\delta}\int_0^\infty e^{-\delta t} \mathbb{E}_z\{R_t(z)\} \mathrm{d}t.$$

We now calculate the average return in our model

$$\mathbb{E}_{z}\left\{R_{t}(z)\right\} = \int R_{t}(z)\mathrm{d}G(z) = \int r^{*}\left(1 + \lambda\left[\frac{z}{\underline{z}(t)} - 1\right]^{+}\right)\eta z^{-\eta-1}\mathrm{d}z = r^{*}\left(1 + \frac{\lambda}{\eta - 1}\underline{z}^{-\eta}\right),$$

where we have used (8) and (10) to express  $R_t(z)$  and integrated using the Pareto productivity distribution. Finally, using (17), we can rewrite:

$$\mathbb{E}_{z}\{R_{t}(z)\} = r^{*} + \frac{\alpha}{\eta} \frac{y(x(t), \ell(t))}{x(t)}$$

which corresponds to equation (20) in the text. Substituting it into (A5) delivers another useful characterization of the value function of entrepreneurs. A similar derivation can be immediately applied to the case with an asset subsidy,  $\varsigma_x(t)$ , as long as it is finite.

## A2 An Economy with Transfers

## A2.1 Desirability of transfers from workers to entrepreneurs

**Proposition A3** Consider a (small) transfer of wealth  $\hat{x}_0 = -\hat{b}_0 > 0$  at t = 0 from a representative household uniformly to all entrepreneurs and a reverse transfer at time t' > 0 equal to

$$\hat{x}_0 \exp\left\{r^*t' + \gamma \int_0^{t'} \frac{\alpha}{\eta} \frac{y(x(t), \ell(t))}{x(t)} \mathrm{d}t\right\} > \hat{x}_0 e^{r^*t'},$$

holding constant  $\ell(t)$  and  $c_e(t)$  for all  $t \ge 0$ . Such perturbation strictly improves the welfare of workers and leaves the welfare of all entrepreneurs unchanged, constituting a Pareto improvement.

**Proof:** For any time path  $\{c, \ell, b, x, c_e\}_{t \ge 0}$  satisfying the household and entrepreneurs budget constraints:

$$\dot{b}(t) = (1 - \alpha)y(x(t), \ell(t)) + r^*b(t) - c(t),$$
(A9)

$$\dot{x}(t) = \frac{\alpha}{\eta} y \big( x(t), \ell(t) \big) + r^* x(t) - c_e(t), \tag{A10}$$

starting from  $(b_0, x_0)$ , consider a perturbation  $\tilde{x}(t) \equiv x(t) + \beta \hat{x}(t)$ , where  $\beta$  is a scalar and  $\hat{x}$  is a differentiable function from  $\mathbb{R}_+$  to  $\mathbb{R}$ , and similarly for other variables. Finally, consider perturbations such that:

$$\begin{aligned} \hat{x}(0) &= -\hat{b}(0) = \hat{x}_0 > 0, \\ \hat{\ell}(t) &= \hat{c}_e(t) = 0 \qquad \forall t \ge 0, \\ \hat{c}(t) &= 0 \qquad \forall t \in [0, t'], \end{aligned}$$

and  $\{\tilde{c}, \tilde{\ell}, \tilde{b}, \tilde{x}, \tilde{c}_e\}_{t \in (0, t')}$  satisfy (A9)–(A10).

For such perturbations, we Taylor-expand (A9)–(A10) around  $\beta = 0$  for  $t \in (0, t')$ :

$$\begin{split} \dot{\hat{b}}(t) &= (1-\alpha) \frac{\partial y\big(x(t),\ell(t)\big)}{\partial x} \hat{x}(t) + r^* \hat{b}(t), \\ \dot{\hat{x}}(t) &= \frac{\alpha}{\eta} \frac{\partial y\big(x(t),\ell(t)\big)}{\partial x} \hat{x}(t) + r^* \hat{b}(t), \end{split}$$

with  $\hat{x}(0) = -\hat{b}(0) = \hat{x}_0$ . Note that these equations are linear in  $\hat{x}(t)$  and  $\hat{b}(t)$ , and we can integrate

them on (0, t) for  $t \leq t'$  to obtain:

$$\hat{b}(t) = -\hat{x}_0 e^{r^* t} + \int_0^t e^{r^* (t-\tilde{t})} (1-\alpha) \frac{\partial y(x(\tilde{t}), \ell(\tilde{t}))}{\partial x} \hat{x}(\tilde{t}) \mathrm{d}\tilde{t},$$
$$\hat{x}(t) = \hat{x}_0 \exp\left\{\int_0^t \left(\frac{\alpha}{\eta} \frac{\partial y(x(\tilde{t}), \ell(\tilde{t}))}{\partial x} + r^*\right) \mathrm{d}\tilde{t}\right\},$$

Therefore, by t = t', we have a cumulative deviation in the state variables equal to:

$$\hat{x}(t'_{-}) + \hat{b}(t'_{-}) = \hat{x}_{0}e^{r^{*}t'} \left[ \left( \exp\left\{ \gamma \int_{0}^{t'} \frac{\alpha}{\eta} \frac{y(x(t), \ell(t))}{x(t)} dt \right\} - 1 \right) + (1 - \gamma) \int_{0}^{t'} e^{-r^{*}t} \frac{\alpha}{\eta} \frac{y(x(t), \ell(t))}{x(t)} \frac{\hat{x}(t)}{\hat{x}_{0}} dt \right],$$

where  $t'_{-}$  denotes an instant before t', and we have used the functional form for  $y(\cdot)$  and definition of  $\gamma$  in (16), which imply  $\partial y/\partial x = \gamma y/x$  and  $(1 - \alpha)\gamma = (1 - \gamma)\alpha/\eta$ . Both terms inside the square bracket are positive (since  $\hat{x}(t)/\hat{x}_0 > 1$  due to the accumulation of the initial transfer). The first term is positive due to the higher return the entrepreneurs make on the initial transfer  $\hat{x}_0$ relative to households. The second term represents the increase in worker wages associated with the higher entrepreneurial wealth, which leads to an improved allocation of resources and higher labor productivity.<sup>2</sup>

At t = t', a reverse transfer from entrepreneurs to workers equal to

$$\hat{x}_0 \exp\left\{r^*t' + \gamma \int_0^{t'} \frac{\alpha}{\eta} \frac{y(x(t), \ell(t))}{x(t)} \mathrm{d}t\right\}$$

results in  $\hat{x}(t') = 0$  and  $\hat{b}(t') > 0$ , which allows to have  $\hat{c}(t) = r^*\hat{b}(t') > 0$  for all  $t \ge t'$ , with  $\hat{\ell}(t) = \hat{c}_e(t) = 0$ . This constitutes a Pareto improvement since the new allocation has the same labor supply by workers and consumption by entrepreneurs with a strictly higher consumption for workers:  $\tilde{\ell}(t) = \ell(t)$ ,  $\tilde{c}_e(t) = c_e(t)$ ,  $\tilde{c}(t) \ge c(t)$  for all  $t \ge 0$  and with strict inequality for  $t \ge t'$ .

#### A2.2 Optimal policy with transfers to entrepreneurs

This appendix shows that the conclusions obtained in Section 3, in particular that optimal Ramsey policy involves a labor subsidy when entrepreneurial wealth is low, are robust to allowing for transfers to entrepreneurs as long as these are constrained to be finite. Formally, we extend the planner's problem (P1) to allow for an asset subsidy to entrepreneurs,  $\varsigma_x$ . In particular, the budget constraints of workers, entrepreneurs, and the government (21), (12) and (22) become

$$c + b \le (1 - \tau_{\ell})w\ell + (r^* - \tau_b)b + T,$$
  
$$\dot{a} = \pi(a, z) + (r^* + \varsigma_x)a - c_e,$$
  
$$\tau_{\ell}w\ell + \tau_b b = \varsigma_x x + T.$$

<sup>2</sup>Note that for small t', we have the following limiting characterization:

$$\frac{\hat{x}(t') + \hat{b}(t')}{\hat{x}_0 t'} \to \frac{\alpha}{\eta} \frac{y(x(0), \ell(0))}{x(0)} \quad \text{as} \quad t' \to 0,$$

which corresponds to the average return differential between entrepreneurs and workers,  $\mathbb{E}_z R_0(z) - r^*$ .

Note that the asset (savings) subsidy to entrepreneurs,  $\varsigma_x x$ , acts as a tool for redistributing wealth from workers to entrepreneurs (or vice versa when  $\varsigma_x < 0$ ). In fact, the asset subsidy is essentially equivalent to a lump-sum transfer to entrepreneurs, as it does not distort the policy functions of either workers or entrepreneurs. The only difference with a lump-sum transfer is that a proportional tax to assets does not affect the consumption policy rule of the entrepreneurs, in contrast to a lumpsum transfer which makes the savings decision of entrepreneurs analytically intractable.<sup>3</sup> In what follows we refer to  $\varsigma_x$  as transfers to entrepreneurs to emphasize that it is a very direct tool for wealth redistribution towards entrepreneurs. Note from (22) that a priori we do not restrict whether it is workers or entrepreneurs who receive revenues from the use of the distortionary taxes  $\tau_\ell$  and  $\tau_b$  (or who pay lump-sum taxes in the case of subsidies).

The planner now chooses a sequence of three taxes,  $\{\tau_b, \tau_\ell, \varsigma_x\}_{t\geq 0}$  to maximize household utility (1) subject to the resulting allocation being a competitive equilibrium. We again make use of Lemma 3, which allows us to recast this problem as the one of choosing a dynamic allocation  $\{c, \ell, b, x\}_{t\geq 0}$  and a sequence of transfers  $\{\varsigma_x\}_{t\geq 0}$  which satisfy the household budget constraint and the aggregate wealth accumulation equation.

We impose an additional constraint on the aggregate transfer:<sup>4</sup>

$$s \le \varsigma_x(t) \, x(t) \le S,\tag{A11}$$

where  $s \leq 0$  and  $S \geq 0$ . Section 3 analyzed the special case of s = S = 0. The case with unrestricted transfers corresponds to  $S = -s = +\infty$ , which we consider as a special case now, but in general we allow s and S to be bounded.

The planning problem for the case with transfers is:

$$\max_{\substack{\{c,\ell,b,x\}_{t\geq 0},\\ \{\varsigma_x: s\leq \varsigma_x x\leq S\}_{t\geq 0}\\ \text{subject to}}} \int_0^\infty e^{-\rho t} u(c,\ell) dt$$
$$(P2)$$
$$\frac{\dot{\zeta}_x}{\dot{\zeta}_x} = \frac{\alpha}{\eta} y(x,\ell) + (r^* + \varsigma_x - \delta)x,$$

given the initial conditions  $b_0$  and  $x_0$ . We still denote the two co-states by  $\mu$  and  $\mu\nu$ . Appendix A3.1 sets up the Hamiltonian for (P2) and provides the full set of equilibrium conditions. In particular, the optimality conditions (27)–(29) still apply, but now with two additional complementary slackness conditions:

$$\nu \ge 1, \quad \varsigma_x x \le S \quad \text{and} \quad \nu \le 1, \quad \varsigma_x x \ge s.$$
 (A12)

This has two immediate implications. First, as before, the planner never distorts the intertemporal margin of workers, that is  $\tau_b \equiv 0$ . Second, whenever the bounds on transfers are slack,  $s < \zeta_x x < S$ , the co-state for the wealth accumulation constraint is unity,  $\nu = 1$ . In particular, this

 $<sup>^{3}</sup>$ The savings rule of entrepreneurs stays unchanged when lump-sum transfers are unanticipated. In this case the savings subsidy and lump-sum transfers are exactly equivalent. However, the assumption of unanticipated lump-sum transfers is unattractive for several reasons.

<sup>&</sup>lt;sup>4</sup>Why transfers may be constrained in reality is discussed in detail in Section A2.3. Given the reasons discussed there, for example political economy considerations limiting aggregate transfers from workers to entrepreneurs, we find a constraint on the aggregate transfer ( $\varsigma_x x$ ) more realistic than one on the subsidy rate ( $\varsigma_x$ ). However, the analysis of the alternative case is almost identical and we leave it out for brevity. In fact, it is straightforward to generalize (A11) to allow s and S to be functions of aggregate wealth, x(t).



Figure A6: Planner's allocation with unlimited transfers

Note: In panel (a), *transfer* refers to the asset (savings) subsidy to entrepreneurs, which equals  $\zeta_x(0) = \infty$  and  $\zeta_x(t) = -r^*$  for all t > 0, financed by a lump-sum tax on workers, and resulting in the path of entrepreneurial wealth x(t) depicted in panel (b); other variables instantaneously reach their steady state values, while labor and savings wedges (taxes) for workers are set to zero.

is always the case when transfers are unbounded,  $S = -s = +\infty$ . Note that  $\nu = 1$  means that the planner's shadow value of wealth, x, equals  $\bar{\mu}$ —the shadow value of extra funds in the household budget constraint. This equalization of marginal values is intuitive given that the planner has access to a transfer between the two groups of agents. From (28) and (30),  $\nu = 1$  immediately implies that the labor supply condition is undistorted, that is  $\tau_{\ell} = 0.5$  This discussion allows us to characterize the planner's allocation when unbounded transfers are available (see illustration in Figure A6):

**Proposition A4** In the presence of unbounded transfers  $(S = -s = +\infty)$ , the planner distorts neither intertemporal consumption choice, nor intratemporal labor supply along the entire transition path:  $\tau_b(t) = \tau_\ell(t) = 0$  for all t. The steady state is achieved in one instant, at t = 0, and the steadystate asset subsidy equals  $\varsigma_x(t) = \bar{\varsigma}_x = -r^*$  for t > 0, i.e. a transfer of funds from entrepreneurs to workers. When  $x(0) < \bar{x}$ , the planner makes an unbounded transfer from workers to entrepreneurs at t = 0, i.e.  $\varsigma_x(0) = +\infty$ , to ensure  $x(0+) = \bar{x}$ .<sup>6</sup>

Proposition A4 shows that the asset subsidy to entrepreneurs dominates the other instruments at the planner's disposal, as long as it is unbounded. When the planner can freely reallocate wealth

$$\dot{m} = (1 - \alpha + \alpha/\eta)y(x,\ell) + r^*m - \delta x - c.$$

<sup>&</sup>lt;sup>5</sup>Note that when transfers are unbounded, (P2) can be replaced with a simpler optimal control problem (P3) with a single state variable  $m \equiv b + x$  and one aggregate dynamic constraint:

The choice of x in this case becomes static, maximizing the right-hand side of the dynamic constraint at each point in time, and the choice of labor supply can be immediately seen to be undistorted. The results of Proposition A4 can be obtained directly from this simplified formulation (see Appendix A3.1).

<sup>&</sup>lt;sup>6</sup>The steady-state entrepreneurial wealth is determined from (26) substituting in  $\bar{\varsigma}_x$ :  $\delta = \alpha/\eta \cdot y(\bar{x}, \ell)/\bar{x}$ , where  $\bar{\ell}$  satisfies the labor supply condition (28) with  $\tau_{\ell} = \gamma(1-\nu) = 0$  and  $u_c = \bar{\mu}$ .



Figure A7: Planner's allocation with limited transfers

Note: Evolution of the labor tax  $\tau_{\ell}(t)$  and entrepreneurial wealth x(t) when transfers to entrepreneurs are bounded by  $\varsigma_x(t)x(t) \leq S < \infty$ , while s is not binding (i.e.,  $s \leq -r^*\bar{x}$ ); until the steady state is reached, the transfer is maxed out,  $\varsigma_x(t) = S/x(t)$ ; in steady state,  $\bar{x}$  and  $\bar{\varsigma}_x = -r^*$  are the same as in Figure A6.

between households and entrepreneurs, she no longer faces the need to distort the labor supply or savings decisions of the workers. Clearly, the infinite transfer in the initial period,  $\varsigma_x(0)$ , is an artifact of the continuous-time environment. In discrete time, the required transfer is simply the difference between initial and steady-state wealth, which however can be very large if the economy starts far below its steady state in terms of entrepreneurial wealth. There is a variety of reasons why large redistributive transfers may be undesirable or infeasible in reality, as we discuss in detail in Appendix A2.3, and alluded to in Section 2.3. We, therefore, turn now to the analysis of the case with bounded transfers.

For brevity, we consider here the case in which the upper bound is binding,  $S < \infty$ , but the lower bound is not binding, that is  $s \leq -r^*\bar{x}$ , while Appendix A3.1 presents the general case. The planner's allocation in this case is characterized by  $u_c = \bar{\mu}$ , (26), (28), (29) and (A12), and the transition dynamics has two phases. In the first phase,  $x(t) < \bar{x}$  and  $\tau_{\ell}(t) < 0$  (as  $\nu(t) > 1$ ), while the planner simultaneously chooses the maximal possible transfer from workers to entrepreneurs each period,  $\varsigma_x(t)x(t) = S$ . During this phase, the characterization is the same as in Proposition 1, but with the difference that a transfer S is added to the entrepreneurs' wealth accumulation constraint (26) and subtracted from the workers' budget constraint (25). That is, starting from  $x_0 < \bar{x}$ , entrepreneurial assets accumulate over time and the planner distorts labor supply upwards at a decreasing rate: x(t) increases and  $\tau_{\ell}(t) < 0$  decreases in absolute value towards zero. The second phase is reached at some finite time  $\bar{t} > 0$ , and corresponds to a steady state described in Proposition A4:  $x(t) = \bar{x}$ ,  $\nu(t) = 1$ ,  $\tau_{\ell}(t) = 0$  and  $\varsigma_x(t) = -r^*$  for all  $t \geq \bar{t}$ . Throughout the entire transition the intertemporal margin of workers is again not distorted,  $\tau_b(t) = 0$  for all t.

We illustrate the planner's dynamic allocation in this case in Figure A7 and summarize its properties in the following Proposition:

**Proposition A5** Consider the case with  $S < \infty$ ,  $s \leq -r^*\bar{x}$ , and  $x(0) < \bar{x}$ . Then there exists  $\bar{t} \in (0, \infty)$  such that: (1) for  $t \in [0, \bar{t})$ ,  $\varsigma_x(t)x(t) = S$  and  $\tau_\ell(t) < 0$ , with the dynamics of  $(x(t), \tau_\ell(t))$  described by a pair of ODEs (26) and (29) together with a static equation (28) (and definition (30)), with a globally stable saddle path as in Proposition 1; (2) for  $t \geq \bar{t}$ ,  $x(t) = \bar{x}$ ,  $\tau_\ell(t) = 0$  and  $\varsigma_x(t) = -r^*$ , corresponding to the steady state in Proposition A4. For all  $t \geq 0$ ,  $\tau_b(t) = 0$ .

Therefore, our main result that optimal Ramsey policy involves a labor supply subsidy when entrepreneurial wealth is low is robust to allowing for transfers from workers to entrepreneurs as long as these transfers are bounded. Applying this logic to a discrete-time environment, whenever the transfers cannot be large enough to jump entrepreneurial wealth immediately to its steadystate level (therefore, resulting in a transition period with  $\nu > 1$ ), the optimal policy involves a pro-business intervention of increasing labor supply.

## A2.3 Infeasibility of transfers

The analysis in Appendix A2.2 suggests the superiority of transfers to alternative policy tools. Here we discuss a number of arguments why transfers may not constitute a feasible or desirable policy option, as well as other constraints on implementation, which justify our focus on the optimal policy under a restricted set of instruments.

First, large transfers may be infeasible simply due to the budget constraint of the government (or the household sector), when the economy starts far away from its long-run level of wealth. Furthermore, unmodeled distributional concerns in a richer environment with heterogeneous workers may make large transfers—which are large lump-sum taxes from the point of view of workers undesirable or infeasible (see Werning, 2007). Note that, in contrast, the policy of subsidizing labor supply, while in the short run also shifting gains towards the entrepreneurial sector, has the additional advantage of increasing GDP and incomes of all groups of agents in the economy. If not just entrepreneurs but also the household sector were financially constrained, or if there were an occupational choice such that workers had the option to become entrepreneurs, large lump-sum taxes on households would be even more problematic and the argument in favor of a labor supply subsidy would be even stronger.

Second, large transfers from workers to entrepreneurs may be infeasible for political economy reasons. This limitation is particularly relevant under socialist or populist governments of many developing countries, but even for more technocratic governments a policy of direct financial injections into the business sector, often labelled as a bailout, may be hard to justify. In contrast, it is probably easier to ensure broad public support of more indirect policies, such as labor supply subsidies or competitive exchange rate devaluations. Another political economy concern is that transfers to businesses may become entrenched once given out, e.g. due to political connections. As a result, originally "well-intended" transfers may persist far beyond what is optimal from the point of view of a benevolent planner (see Buera, Moll, and Shin, 2013).

Third, the information requirement associated with transfers is likely to be unrealistically strict. Indeed, the government needs to be able to separate entrepreneurs from workers, as every individual in the economy will have an incentive to declare himself an entrepreneur when the government announces the policy of direct subsidies to business. As a result, the government is likely to be forced to condition its support on some easily verifiable observables. One potential observable is the amount of labor hired by entrepreneurs, and the labor supply subsidy implicitly does just that.<sup>7</sup>

<sup>&</sup>lt;sup>7</sup>For tractability, the way we set up the Ramsey problem without transfers, the subsidy to labor supply is

Furthermore, and as already mentioned in Section 2.3, transfers constitute such a powerful tool in our environment because they allow the government to effectively side-step the collateral constraint in the economy, by first inflating entrepreneurial wealth and later imposing a lump-sum tax on entrepreneurs to transfer the resources back to the households. Such a policy may be infeasible if entrepreneurs can hide their wealth from the government. In contrast, labor supply taxes are less direct, affecting entrepreneurs only via the equilibrium wage rate, and hence less likely to trigger such deviations.

Finally, the general lesson from our analysis is the optimality of a pro-business stance of government policy during the initial phase of the transition, which may be achieved to some extent with whatever instrument the government has at its disposal. It is possible that the government has very limited flexibility in the use of any tax instruments, and hence has to rely on alternative non-tax market regulation. For example, the government can choose how much market and bargaining power to leave to each group of agents in the economy, or affect the market outcomes by means of changing the value of the outside options of different agents.<sup>8</sup> Such interventions may allow the government to implement some of the Ramsey-optimal allocations without the use of explicit taxes and transfers.

## A3 Derivations and Proofs for Section 3

## A3.1 Optimality conditions for the planner's problem

Consider the generalization of the planner's problem (P1), which allows for (possibly bounded) direct transfers between workers and entrepreneurs, as we set it up in (P2) in Appendix A2.2. Without loss of generality, we normalize these transfers to be in proportion with entrepreneurial wealth,  $\varsigma_x x$ , and denote with s and S the lower (possibly negative) and upper bounds on these transfers respectively. With the transfers, the constraints on the planner's problem (25) and (26) are simply adjusted by quantity  $\varsigma_x x$ , with a negative sign in the first case and a positive sign in the second. The present-value Hamiltonian associated with the planner's problem (P2) is:

$$\mathcal{H} = u(c,\ell) + \mu \left[ (1-\alpha)y(x,\ell) + r^*b - c - \varsigma_x x \right] + \mu \nu \left[ \frac{\alpha}{\eta} y(x,\ell) + (r^* + \varsigma_x - \delta)x \right] + \mu \bar{\xi}(S - \varsigma_x x) + \mu \underline{\xi}(\varsigma_x x - s),$$

where we have introduced two additional Lagrange multipliers  $\mu \bar{\xi}$  and  $\mu \bar{\xi}$  for the corresponding bounds on transfers in (A11). The full set of optimality conditions is given by:

financed by a lump-sum tax on workers. An alternative formulation is to levy the lump-sum tax on all agents in the economy without discrimination. The two formulations yield identical results in the limiting case when the number (mass) of entrepreneurs is diminishingly small relative to the number (mass) of workers.

<sup>&</sup>lt;sup>8</sup>During the New Deal policies of Franklin D. Roosevelt, the government increased the monopoly power of unions in the labor market and businesses in the product markets (see Cole and Ohanian, 2004, for a quantitative analysis of these policies in the context of a neoclassical growth model). Many Asian countries, for example Korea, have taken an alternative pro-business stance in the labor market, by halting unions and giving businesses an effective monopsony power. The governments of relatively rich European countries, on the other hand, tilt the bargaining power in favor of labor by providing generous unemployment insurance and a strict regulation of hiring and firing practices. See Online Appendix B for a historical account of various tax and non-tax market regulation policies adopted across a number of countries.

$$0 = \frac{\partial \mathcal{H}}{\partial c} = u_c - \mu, \tag{A13}$$

$$0 = \frac{\partial \mathcal{H}}{\partial \ell} = -u_{\ell} + \mu \left(1 - \gamma + \gamma \nu\right) (1 - \alpha) \frac{y}{\ell},\tag{A14}$$

$$0 = \frac{\partial \mathcal{H}}{\partial \varsigma_x} = \mu x \left( \nu - 1 - \bar{\xi} + \underline{\xi} \right), \tag{A15}$$

$$\dot{\mu} - \rho \mu = -\frac{\partial \mathcal{H}}{\partial b} = -\mu r^*, \tag{A16}$$

$$\dot{(\mu\nu)} - \rho\mu\nu = -\frac{\partial\mathcal{H}}{\partial x} = -\mu(1 - \gamma + \gamma\nu)\frac{\alpha}{\eta}\frac{y}{x} - \mu\nu(r^* - \delta) - \mu\varsigma_x\left(\nu - 1 - \bar{\xi} + \underline{\xi}\right), \quad (A17)$$

where we have used the fact that  $\partial y/\partial \ell = (1 - \gamma)y/\ell$  and  $\partial y/\partial x = \gamma y/x$  which follows from the definition of  $y(\cdot)$  in (16). Additionally, we have two complementary slackness conditions for the bounds-on-transfers constraints:

$$\bar{\xi} \ge 0, \quad \varsigma_x x \le S \quad \text{and} \quad \underline{\xi} > 0, \quad \varsigma_x x \ge s.$$
 (A18)

Under our parameter restriction  $\rho = r^*$ , (A16) and (A13) imply:

$$\dot{\mu} = 0 \qquad \Rightarrow \qquad u_c(t) = \mu(t) \equiv \bar{\mu} \quad \forall t.$$

With this, (A14) becomes (28) in the text. Given  $\mu \equiv \bar{\mu}$ ,  $r^* = \rho$  and (A15), (A17) becomes (29) in the text. Finally, (A15) can be rewritten as:

$$\nu - 1 = \bar{\xi} - \underline{\xi}$$

When both bounds are slack, (A18) implies  $\bar{\xi} = \underline{\xi} = 0$ , and therefore  $\nu = 1$ . When the upper bound is binding,  $\nu - 1 = \bar{\xi} > 0$ , and when the lower bound is binding  $\nu - 1 = -\underline{\xi} < 0$ . Therefore, we obtain the complementary slackness condition (A12).

The case with no transfers (S = -s = 0) results in the planner's problem (P1) with an associated Hamiltonian:

$$\mathcal{H} = u(c,\ell) + \mu \left[ (1-\alpha)y(x,\ell) + r^*b - c \right] + \mu \nu \left[ \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x \right].$$

The optimality conditions in this case are (A13), (A14), (A16) and

$$\dot{(\mu\nu)} - \rho\mu\nu = -\frac{\partial\mathcal{H}}{\partial x} = -\mu(1-\gamma+\gamma\nu)\frac{\alpha}{\eta}\frac{y}{x} - \mu\nu(r^*-\delta),$$

which result in (27)–(29) after simplification.

The case with unbounded transfers  $(S = -s = +\infty)$  allows to simplify the problem considerably, as we discuss in more detail above in Appendix A2.2. Indeed, in this case we can define a single state variable  $m \equiv b + x$ , and sum the two constraints in problem (P2), to write the

resulting problem as:

$$\max_{\substack{\{c,\ell,x,m\}_{t\geq 0}}} \int_0^\infty e^{-\rho t} u(c,1-\ell) dt$$
subject to
$$\dot{m} = (1-\alpha+\alpha/\eta) y(x,\ell) + r^*m - \delta x - c,$$
(P3)

with a corresponding present-value Hamiltonian:

$$\mathcal{H} = u(c, 1-\ell) + \mu \big[ \big(1-\alpha+\alpha/\eta\big) y(x,\ell) + r^*m - \delta x - c \big],$$

with the optimality conditions given by (A13), (A16) and

$$0 = \frac{\partial \mathcal{H}}{\partial \ell} = -u_{\ell} + \mu (1 - \alpha) \frac{y}{\ell}, \tag{A19}$$

$$0 = \frac{\partial \mathcal{H}}{\partial x} = \mu \left( -\delta + \frac{\alpha}{\eta} \frac{y}{x} \right). \tag{A20}$$

(A19) immediately implies  $\tau_{\ell}(t) \equiv 0$ , and (A20) pins down  $x/\ell$  at each instant. The required transfer is then backed out from the aggregate entrepreneurial wealth dynamics (26).

The case with bounded transfers Consider the case with  $S < \infty$ . There are two possibilities: (a)  $s \leq -r^* \bar{x}$ ; and (b)  $r^* \bar{x} < s \leq 0$ , which we consider first. In this case there are two regions:

- 1. for  $x < \bar{x}$ ,  $\varsigma_x x = S$  binds,  $\bar{\xi} = \nu 1 > 0$  and  $\underline{\xi} = 0$ . This immediately implies  $\tau_{\ell} = \gamma(1-\nu) < 0$ , and the dynamics of  $(x, \tau_{\ell})$  are as in Proposition 1, with the difference that  $\dot{x} = \alpha y/\eta + (r^* \delta)x + S$  with S > 0 rather than S = 0.
- 2. when  $x = \bar{x}$  is reached, the economy switches to the steady-state regime with  $\bar{\varsigma}_x \bar{x} = s < 0$  binding, and hence  $\nu 1 = -\xi < 0$  and  $\bar{\xi} = 0$ , in which:

$$\frac{\alpha}{\eta} \frac{y(\bar{x},\ell)}{\bar{x}} = (\delta - r^*) - \frac{s}{\bar{x}} < \delta,$$
$$\bar{\tau}_{\ell} = \gamma \left(1 - \bar{\nu}\right) = \frac{\gamma}{\gamma + (1 - \gamma) \frac{\delta \bar{x}}{r^* \bar{x} + s}} > 0.$$

When the steady-state regime is reached, there is a jump from a labor supply subsidy to a labor supply tax, as well as a switch in the aggregate transfer to entrepreneurs from S to s.

In the alternative case when  $s < -r^*\bar{x}$ , the first region is the same, and in steady state  $\bar{\varsigma}_x\bar{x} = -r^*\bar{x} > s$  and hence the constraint is not binding:  $\underline{\xi} = \overline{\xi} = \overline{\nu} - 1 = \overline{\tau}_{\ell} = 0$ . The steady state in this case is characterized by (A19)–(A20), and  $\bar{\varsigma}_x = -r^*$  ensures  $\dot{x} = 0$  at  $\bar{x}$ . In this case,  $\tau_{\ell}$  continuously increases to zero when steady state is reached, and the aggregate transfer to entrepreneurs jumps from S to  $-r^*\bar{x}$ .

## A3.2 Proof of Proposition 1

Consider (26)–(29). Under our parameter restriction  $\rho = r^*$ , the households' marginal utility is constant over time  $\mu(t) = u_c(t) = \bar{\mu}$  for all t. Using the definition of the Frisch labor supply

function (see Appendix A1.1), (18), (16) and (30), (28) can be written as

$$\ell = \ell^F \Big( \bar{\mu}, (1 - \tau_\ell) (1 - \alpha) \Theta(x/\ell)^\gamma \Big).$$

For given  $(\bar{\mu}, \tau_{\ell}, x)$ , this is a fixed point problem in  $\ell$ , and given positive and finite Frisch elasticity (A3) (i.e., under the condition on the utility function (A4)) one can show that it has a unique solution, which we denote by  $\ell = \ell(x, \tau_{\ell})$ , where we suppress the dependence on  $\bar{\mu}$  for notational simplicity. Note that

$$\frac{\partial \log \ell(x,\tau_{\ell})}{\partial \log x} = \frac{\varepsilon \gamma}{1+\varepsilon \gamma} \in (0,1), \qquad \frac{\partial \log \ell(x,\tau_{\ell})}{\partial \log(1-\tau_{\ell})} = \frac{\varepsilon}{1+\varepsilon \gamma} \in (0,1/\gamma), \tag{A21}$$

where  $\varepsilon$  is the Frisch elasticity defined in (A3), which also implies the bounds in (A21). Substituting  $\ell(x, \tau_{\ell})$  into (29) and (26), we have a system of two autonomous ODEs in  $(\tau_{\ell}, x)$ 

$$\begin{split} \dot{\tau}_{\ell} &= \delta(\tau_{\ell} - \gamma) + \gamma \left(1 - \tau_{\ell}\right) \frac{\alpha}{\eta} \Theta\left(\frac{\ell(x, \tau_{\ell})}{x}\right)^{1 - \gamma}, \\ \dot{x} &= \frac{\alpha}{\eta} \Theta x^{\gamma} \ell(x, \tau_{\ell})^{1 - \gamma} + (r^* - \delta)x. \end{split}$$

We now show that the dynamics of this system in  $(\tau_{\ell}, x)$  space can be described with the phase diagram in Figure 1.

**Steady State** We first show that there exists a unique positive steady state  $(\bar{\tau}_{\ell}, \bar{x})$ , i.e. a solution to

$$\gamma \left(1 - \bar{\tau}_{\ell}\right) \frac{\alpha}{\eta} \Theta \left(\frac{\ell(\bar{x}, \bar{\tau}_{\ell})}{\bar{x}}\right)^{1 - \gamma} = \delta(\gamma - \bar{\tau}_{\ell}), \tag{A22}$$

$$\frac{\alpha}{\eta} \Theta \left( \frac{\ell(\bar{x}, \bar{\tau}_{\ell})}{\bar{x}} \right)^{1-\gamma} = \delta - r^*.$$
(A23)

Substituting (A23) into (A22) and rearranging, we obtain the expression for  $\bar{\tau}_{\ell}$  in (31). From (A23),  $\bar{x}$  is then the solution to the fixed point problem

$$\bar{x} = \left(\frac{\alpha}{\eta} \frac{\Theta}{\delta - r^*}\right)^{\frac{1}{1 - \gamma}} \ell(\bar{x}, \bar{\tau}_\ell) \equiv \Phi(\bar{x}).$$
(A24)

Depending on the properties of the Frisch labor supply function, there may be a trivial solution  $\bar{x} = 0$ . We instead focus on positive steady states. Consider  $\varepsilon(\mu, w)$  from (A3) and define

$$\varepsilon_1 \equiv \min_w \varepsilon(\bar{\mu}, w) > 0, \quad \varepsilon_2 \equiv \max_w \varepsilon(\bar{\mu}, w) < \infty, \quad \theta_1 \equiv \frac{\varepsilon_1 \gamma}{1 + \varepsilon_1 \gamma} > 0, \quad \theta_2 \equiv \frac{\varepsilon_2 \gamma}{1 + \varepsilon_2 \gamma} < 1.$$

From (A21), there are constants  $k_1$  and  $k_2$  such that  $k_1 x^{\theta_1} \leq \ell(x, \bar{\tau}_\ell) \leq k_2 x^{\theta_2}$ . Since  $\theta_1 > 0, \theta_2 < 1$ , there are  $x_1 > 0$  sufficiently small and  $x_2 < \infty$  sufficiently large such that  $\Phi(x_1) > x_1$  and  $\Phi(x_2) < x_2$ . Finally, taking logs on both sides of (A24), we have

$$\tilde{x} = \tilde{\Theta} + \tilde{\ell}(\tilde{x}), \quad \tilde{\ell}(\tilde{x}) \equiv \log \ell(\exp(\tilde{x}), \bar{\tau}_{\ell}), \quad \tilde{\Theta} \equiv \log \left(\frac{\alpha}{\eta} \frac{\Theta}{\delta - r^*}\right)^{\frac{1}{1 - \gamma}},$$
(A25)

satisfying  $\tilde{\Theta} + \tilde{\ell}(\tilde{x}_1) > \tilde{x}_1$  and  $\tilde{\Theta} + \tilde{\ell}(\tilde{x}_2) < \tilde{x}_2$ , where  $\tilde{x}_j \equiv \log x_j$ , for  $j \in \{1, 2\}$ . From (A21), we have  $0 < \tilde{\ell}'(\tilde{x}) < 1$  for all  $\tilde{x}$  and therefore (A25) has a unique fixed point  $\tilde{x}_1 < \log \bar{x} < \tilde{x}_2$ .

**Transition dynamics** (A23) implicitly defines a function  $x = \phi(\tau_{\ell})$ , which is the  $\dot{x} = 0$  locus. We have that

$$\frac{\partial \log \phi(\tau_{\ell})}{\partial \log(1 - \tau_{\ell})} = \frac{\frac{\partial \log \ell}{\partial \log(1 - \tau_{\ell})}}{1 - \frac{\partial \log \ell}{\partial \log x}} = \varepsilon \in (0, \infty).$$

Therefore the  $\dot{x} = 0$  locus is strictly downward-sloping in  $(x, \tau_{\ell})$  space, as drawn in Figure 1. The  $\dot{\tau}_{\ell} = 0$  locus may be non-monotonic, but we know that the two loci intersect only once (the steady state is unique). The state space can then be divided into four quadrants. It is easy to see that  $\dot{\tau}_{\ell} > 0$  for all points to the north-west of the  $\dot{\tau}_{\ell} = 0$  locus, and  $\dot{x} > 0$  for all points to the south-west of the  $\dot{\tau}_{\ell} = 0$  locus, and  $\dot{x} > 0$  for all points to the south-west of the  $\dot{x} = 0$  locus, as indicated by the arrows in Figure 1. It then follows that the system is saddle path stable. Assuming Inada conditions on the utility function and given output function  $y(\cdot)$  defined in (16), the saddle path is the unique solution to the planner's problem (P1).

Now consider points  $(x, \tau_{\ell})$  along the saddle path. There is a threshold  $\hat{x}$  such that  $\tau_{\ell} < 0$  whenever  $x < \hat{x}$  and vice versa, that is labor supply is subsidized when wealth is sufficiently low. There is an alternative argument for this result along the lines of footnote 20 in the text. Equation (29) can be solved forward to yield:

$$\nu(0) = \int_0^\infty e^{-\int_0^t (\delta - \alpha y_x(s)/\eta) \mathrm{d}s} (1 - \alpha) y_x(t) \mathrm{d}t,$$

with  $x(0) = x_0$  and where  $y_x(t) \equiv \partial y(x(t), \ell(t)) / \partial x = \gamma y(x(t), \ell(t)) / x(t) \propto (\ell(t)/x(t))^{1-\gamma}$ . The marginal product of  $x, y_x$ , is unbounded as  $x \to 0$ . Therefore, for low enough  $x_0$ , we must have  $\nu(0) > 1$  and hence  $\tau_\ell(0) < 0$ .

#### A3.3 Additional tax instruments

Consider a planner endowed with the following additional subsidies to entrepreneurs: an asset subsidy  $\varsigma_x$ , a profit subsidy  $\varsigma_{\pi}$ , a sales (revenue) subsidy  $\varsigma_y$ , a capital subsidy  $\varsigma_k$ , and a wagebill subsidy  $\varsigma_w$ . Under these circumstances, the budget set of an entrepreneur can be represented as:

$$\dot{a} = (1 + \varsigma_{\pi})\pi(a, z) + (r^* + \varsigma_x - \delta)a,$$
with
$$\pi(a, z) = \max_{n \ge 0, 0 \le k \le \lambda a} \left\{ (1 + \varsigma_y) A(zk)^{\alpha} n^{1-\alpha} - (1 - \varsigma_w) wn - (1 - \varsigma_k) r^* k \right\},$$
(A26)

which generalizes expression (33) in the text, and where we already incorporated the optimal consumption-saving decision of entrepreneurs, which is  $c_e = \delta a$  independently of the adopted policy instruments.

We next prove an equilibrium characterization result for this case, analogous to Lemma 2:

**Lemma A5** When subsidies  $(\varsigma_x, \varsigma_\pi, \varsigma_y, \varsigma_k, \varsigma_w)$  are used, the output function is given by:

$$y = \left(\frac{1+\varsigma_y}{1-\varsigma_k}\right)^{\gamma(\eta-1)} \Theta x^{\gamma} \ell^{1-\gamma}, \tag{A27}$$

where  $\Theta$  and  $\gamma$  are defined as in Lemma 2, and we have:

$$\underline{z}^{\eta} = \frac{1-\varsigma_k}{1+\varsigma_y} \frac{\eta\lambda}{\eta-1} \frac{r^*}{\alpha} \frac{x}{y},$$
$$(1-\varsigma_w)w\ell = (1-\alpha)(1+\varsigma_y)y,$$
$$(1-\varsigma_k)r^*\kappa = \frac{\eta-1}{\eta}\alpha(1+\varsigma_y)y,$$
$$\Pi = \frac{\alpha}{\eta}(1+\varsigma_y)y.$$

**Proof**: Consider the profit maximization problem (A26). The solution to this problem is given by:

$$k = \lambda a \mathbf{1}_{\{z \ge \underline{z}\}},$$
  

$$n = \left( (1 - \alpha) \frac{(1 + \varsigma^y) A}{(1 - \varsigma^w) w} \right)^{1/\alpha} zk,$$
  

$$\pi = \left[ \frac{z}{\underline{z}} - 1 \right] (1 - \varsigma^k) r^* k,$$

where the cutoff is defined by the zero-profit condition:

$$\alpha \left[ (1+\varsigma^y)A \right]^{1/\alpha} \left( \frac{1-\alpha}{(1-\varsigma^w)w} \right)^{\frac{1-\alpha}{\alpha}} \underline{z} = (1-\varsigma^k)r^*.$$
(A28)

Finally, labor demand in the sector is given by:

$$\ell = \left( (1-\alpha) \frac{(1+\varsigma^y)A}{(1-\varsigma^w)w} \right)^{1/\alpha} \frac{\eta \lambda}{\eta - 1} x \underline{z}^{1-\eta},$$
(A29)

and aggregate output is given by:

$$y = \left( (1-\alpha)\frac{(1+\varsigma^y)}{(1-\varsigma^w)w} \right)^{\frac{1-\alpha}{\alpha}} A^{1/\alpha} \frac{\eta\lambda}{\eta-1} x \underline{z}^{1-\eta}.$$
 (A30)

Combining these three conditions, we solve for  $\underline{z}$ , w and y, which result in the first three equations of the lemma. Aggregate capital demand and profits in this case are given by:

$$\kappa = \lambda x \underline{z}^{-\eta}$$
 and  $\Pi = (1 - \varsigma^k) r^* \kappa / (\eta - 1),$ 

and combining these with the solution for  $\underline{z}^{\eta}$  we obtain the last two equations of the lemma.

The immediate implication of this lemma is that asset and profit subsidies do not affect the equilibrium relationships directly, but do so only indirectly through their effect on aggregate entrepreneurial wealth.

With this characterization, and given that the subsidies are financed by a lump-sum tax on households, we can write the planner's problem as

$$\max_{\{c,\ell,b,x,\varsigma_x,\varsigma_\pi,\varsigma_k,\varsigma_w,\varsigma_y\}} \int_0^\infty e^{-\rho t} u\big(c(t),\ell(t)\big) \mathrm{d}t \tag{P4}$$

subject to

$$c+\dot{b} \leq \left[ (1-\alpha) - \frac{\varsigma_y}{1+\varsigma_y} - \frac{\varsigma_k}{1-\varsigma_k} \frac{\eta-1}{\eta} \alpha - \varsigma_\pi \frac{\alpha}{\eta} \right] (1+\varsigma_y) y(x,\ell,\varsigma_y,\varsigma_k) + r^* b - \varsigma_x x,$$
  
$$\dot{x} = (1+\varsigma_\pi) \frac{\alpha}{\eta} (1+\varsigma_y) y(x,\ell,\varsigma_y,\varsigma_k) + (r^*+\varsigma_x-\delta) x,$$

where  $y(x, \ell, \varsigma_y, \varsigma_k)$  is defined in (A27) and the negative terms in the square brackets correspond to lump-sum taxes levied to finance the respective subsidies. Note that  $\varsigma_w$  drops out from the constraints (just like w does in Lemma 2). It can be recovered from

$$-\frac{u_c}{u_\ell} = (1 - \tau_\ell)w = \frac{1 - \tau_\ell}{1 - \varsigma_w} \cdot (1 + \varsigma_y)(1 - \alpha)\frac{y}{\ell},$$
(A31)

under the additional assumption  $\tau_{\ell} = 0$ . Without this assumption there is implementational indeterminacy since  $\tau_{\ell}$  and  $\varsigma_w$  are perfectly substitutable policy instruments as long as  $(1 - \tau_{\ell})/(1 - \varsigma_w)$ remains constant.

When unbounded asset or profit subsidies are available, we can aggregate the two constraints in (P4) in the same way we did in the planner's problem (P3) in Appendix A3.1 by defining a single state variable  $m \equiv b + x$ . The corresponding Hamiltonian in this case is:

$$\mathcal{H} = u(c,\ell) + \mu \left[ \left( 1 - \alpha + \frac{\alpha}{\eta} - \frac{\varsigma_y}{1 + \varsigma_y} - \frac{\varsigma_k}{1 - \varsigma_k} \frac{\eta - 1}{\eta} \alpha \right) \frac{(1 + \varsigma_y)^{1 + \gamma(\eta - 1)}}{(1 - \varsigma_k)^{\gamma(\eta - 1)}} \Theta x^{\gamma} \ell^{1 - \gamma} + r^* m - \delta x - c \right],$$

where we have substituted (A27) for y. The optimality with respect to  $(\varsigma_y, \varsigma_k)$  evaluated at  $\varsigma_y = \varsigma_k = 0$  are, after simplification:

$$\frac{\partial \mathcal{H}}{\partial \varsigma_y}\Big|_{\varsigma_y = \varsigma_k = 0} \propto -\frac{1}{1 - \alpha + \alpha/\eta} + 1 + \gamma(\eta - 1) = 0,$$
  
$$\frac{\partial \mathcal{H}}{\partial \varsigma_k}\Big|_{\varsigma_y = \varsigma_k = 0} \propto -\frac{\frac{\eta - 1}{\eta}\alpha}{1 - \alpha + \alpha/\eta} + \gamma(\eta - 1) = 0,$$

and combining  $\partial \mathcal{H}/\partial c = 0$  and  $\partial \mathcal{H}/\partial \ell = 0$ , both evaluated at  $\varsigma_y = \varsigma_k = 0$ , we have:

$$-u_{\ell}/u_c = (1-\alpha)y/\ell.$$

Finally, optimality with respect to m implies as before  $\dot{\mu} = 0$  and  $u_c(t) = \mu(t) \equiv \bar{\mu}$  for all t. This implies that whenever profit and/or asset subsidies are available and unbounded, other instruments are not used:

$$\varsigma_y = \varsigma_k = \varsigma_w - \tau_\ell = \varsigma_b = 0.$$

Indeed, both  $\varsigma_{\pi}$  and  $\varsigma_x$ , appropriately chosen, act as transfers between workers and entrepreneurs, and do not affect any equilibrium choices directly, in particular do not affect  $y(\cdot)$ , as can be seen from (A27). This is the reason why these instruments are favored over other distortionary ways to affect the dynamics of entrepreneurial wealth, just like in Propostion A4 in Appendix A2.2.

Examining (33), we see that the following combination of taxes  $\varsigma_y = -\varsigma_k = -\varsigma_w = \varsigma$  is equivalent to a profit subsidy  $\varsigma_{\pi} = \varsigma$ , and therefore whenever these three instruments are jointly available, they are used in this way to replicate a profit subsidy.

Next, in the planner's problem (P4) we restrict  $\varsigma^x = \varsigma^\pi \equiv 0$ , and write the resulting Hamilto-

nian:

$$\mathcal{H} = u(c,\ell) + \mu \left[ r^* b - c + \left( (1-\alpha) - \frac{\varsigma^y}{1+\varsigma^y} - \frac{\varsigma^k}{1-\varsigma^k} \frac{\eta-1}{\eta} \alpha \right) (1+\varsigma^y) y \right] + \mu \nu \left[ (r^* - \delta)x + \frac{\alpha}{\eta} (1+\varsigma^y) y \right],$$

where y is given in (A27). The optimality conditions with respect to b and c are as before, and result in  $u_c = \mu \equiv \overline{\mu}$ . Optimality with respect to x results in a dynamic equation for  $\nu$ , analogous to (29). And the optimality conditions with respect to  $\varsigma^k$ ,  $\varsigma^y$  and  $\ell$  are now given by:

$$\begin{split} \frac{\partial \mathcal{H}}{\partial \varsigma^k} &\propto -\left[\frac{\varsigma^y}{1+\varsigma^y} + \frac{\varsigma^k}{1-\varsigma^k}\right] + \frac{\alpha}{\eta}(\nu-1) = 0,\\ \frac{\partial \mathcal{H}}{\partial \varsigma^y} &\propto -(\eta-1)\left[\frac{\varsigma^y}{1+\varsigma^y} + \frac{\varsigma^k}{1-\varsigma^k}\right] + (\nu-1) = 0,\\ \frac{\partial \mathcal{H}}{\partial \ell} &\propto \frac{u_\ell}{u_c} + \left(1 - \gamma \frac{\eta}{\alpha} \frac{\varsigma^y}{1+\varsigma^y} - \gamma(\eta-1) \frac{\varsigma^k}{1-\varsigma^k} + \gamma(\nu-1)\right) \frac{(1+\varsigma^y)(1-\alpha)y}{\ell} = 0. \end{split}$$

We consider the case when there is an additional restriction—either  $\varsigma^y = 0$  or  $\varsigma^k = 0$ —so that a profit subsidy cannot be engineered. We immediately see that in the former case we obtain (34), which proves the claim in Proposition 2.<sup>9</sup>

**Proof of Proposition 2** Consider the optimality conditions above after imposing  $\varsigma^y = 0$ . The first of them immediately implies:

$$\frac{\varsigma^k}{1-\varsigma^k} = \frac{\alpha}{\eta}(\nu-1).$$

The second of them does not hold, because  $\varsigma^y = 0$  rather than chosen optimally. Finally, manipulating the third one, we get:

$$\begin{aligned} -\frac{u_{\ell}}{u_c} &= \left(1 - \gamma(\eta - 1)\frac{\varsigma^k}{1 - \varsigma^k} + \gamma(\nu - 1)\right)\frac{(1 - \alpha)y}{\ell} \\ &= \left(1 + \frac{\alpha}{\eta}(\nu - 1)\right)\frac{(1 - \alpha)y}{\ell}. \end{aligned}$$

where the second line substitutes in the expression for the optimal  $\varsigma^k$ . This last expression characterizes the optimal labor wedge, so that from (A31) we have:

$$\frac{1-\tau_{\ell}}{1-\varsigma_w} = 1 + \frac{\alpha}{\eta}(\nu-1) \qquad \xrightarrow{\tau_{\ell}=0} \qquad \frac{\varsigma_w}{1-\varsigma_w} = \frac{\alpha}{\eta}(\nu-1),$$

completing the proof of the claim in the proposition.

<sup>9</sup>In the alternative case with  $\varsigma^k = 0$ , the optimal use of the sales and wagebill subsidies is characterized by:

$$\frac{\varsigma^y}{1+\varsigma^y} = -\frac{\varsigma^w}{1-\varsigma^w} = \frac{\nu-1}{\eta-1}$$

with the overall labor wedge  $\varsigma \equiv \frac{1+\varsigma^y}{1-\varsigma^w} - 1 = 0$ . That is, if both a revenue and a labor subsidy are present, a pro-business policy can be implemented without a labor wedge, but this nonetheless requires the use of the labor tax to partly offset the distortion created by the sales subsidy.

#### A3.4 Finite lives and financially constrained households

Alternative social welfare and time inconsistency As an alternative to (35), consider:

$$\tilde{W}_{0} = \int_{0}^{\infty} e^{-\varrho\tau} q U_{0}(\tau) \mathrm{d}\tau + \int_{-\infty}^{0} q e^{q\tau} \int_{0}^{\infty} e^{-(\rho+q)t} u_{\tau}(t) \mathrm{d}t \mathrm{d}\tau.$$
(A32)

This criterion discounts remaining lifetime utility of those currently alive ( $\tau \leq 0$ ) to time t = 0rather than to their birth at time  $t = \tau < 0$ , or in other words the planner uses the remaining lifetime utility  $\tilde{U}_0(\tau) = \int_0^\infty e^{-(\rho+q)t} u_\tau(t) dt$  for  $\tau < 0$ . Then the planner integrates remaining lifetime utilities across the living ( $\tau < 0$ ) using their population density  $qe^{q\tau}$  and adds the plannerdiscounted future lifetime utilities of the unborn ( $\tau \geq 0$ ). When  $\varrho = \rho$ , the alternative welfare criterion (A32) is equivalent to (35). However, when  $\varrho \neq \rho$ , they are not, and the alternative criterion in (A32) causes a time inconsistency problem for the planner. Indeed, at time t = 0, she discounts heavily (assuming  $\varrho > \rho$ ) the unborn future cohorts, but as they are being born, their weight in the social welfare increases, making the planner want to deviate from the earlier plan. This problem is avoided with the social welfare function in (35), which we adopt for our analysis, and which maintains consistency in the planner's weights on different cohorts at different points in time.

**Optimality conditions with present bias and borrowing constraints** The planner's problem is now:

$$\max_{\{c,\ell,b,x\}_{t\geq 0}} \int_0^\infty e^{-\varrho t} u(c_t,\ell_t) \mathrm{d}t,$$

where  $\rho \leq \rho \leq \rho + q$  and  $\rho = r^*$ , and subject to:

$$\dot{b} = r^*b + (1 - \alpha)y(x, \ell) - c$$
$$\dot{x} = \frac{\alpha}{n}y(x, \ell) + (r^* - \delta)x,$$

where as before  $y(x, \ell) = \Theta x^{\gamma} \ell^{1-\gamma}$ , and in addition possibly subject to  $b \ge 0$  (with initial condition  $b_0 = 0$  and  $x_0 > 0$ ). The associated present-value Hamiltonian is:

$$\mathcal{H} = u(c,\ell) + \mu \left[ r^* b + (1-\alpha)y(x,\ell) - c \right] + \mu \nu \left[ \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x \right] + \iota_b \psi b,$$

where  $\iota_b \in \{0, 1\}$  for whether the borrowing constraint is imposed on the households (and the planner). The optimality conditions are:

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial c} &= u_c - \mu = 0, \\ \frac{\partial \mathcal{H}}{\partial \ell} &= u_\ell + \mu \left[ (1 - \alpha) + \nu \frac{\alpha}{\eta} \right] y_\ell = 0, \\ \dot{\mu} - \varrho \mu &= -\frac{\partial \mathcal{H}}{\partial b} = -\mu r^* - \iota_b \psi, \\ \dot{(\mu\nu)} - \varrho \mu \nu &= -\frac{\partial \mathcal{H}}{\partial x} = -\mu \left[ (1 - \alpha) + \nu \frac{\alpha}{\eta} \right] y_x - \mu \nu (r^* - \delta). \end{aligned}$$

We have  $u_c = \mu$ , and given  $\rho = r^*$ , we rewrite the optimality for b as:

$$\frac{\mu}{\mu} = \underbrace{\varrho - \rho}_{\geq 0} - \iota_b \psi$$

The remaining two conditions characterize the optimal labor wedge:

$$-\frac{u_{\ell}}{u_{c}} = \left[1 + \overbrace{\gamma(\nu-1)}^{=\tau_{\ell}}\right](1-\alpha)\frac{y}{\ell},$$
$$\underbrace{\nu - \underbrace{\left(\varrho - \rho + \delta - \frac{\dot{\mu}}{\mu}\right)}_{\ell}\nu = -\left[1 + \gamma(\nu-1)\right]\frac{\alpha}{\eta}\frac{y}{x}.$$

First, consider the case without borrowing constraints on households (planner), so that  $\iota_b = 0$ . Then the two optimality conditions are identical to those in the baseline model, (28)–(29). Furthermore, the resulting policy is exactly the same as in the baseline model conditional on the path of labor supply  $\{\ell\}$ , and the only difference in the path of the planner's allocation may arise due to the income effect of  $\{c\}$  on labor supply  $\{\ell\}$ . Indeed, a planner with  $\rho > \rho$  chooses a declining path of consumption since  $u_c = \mu$  and  $\dot{\mu}/\mu = \rho - \rho > 0$  (i.e., front-loading of consumption to earlier generations by means of international borrowing), in contrast with a flat consumption profile in the baseline model ( $u_c = \bar{\mu} = const$ ). However, if the preferences are GHH with no income effect on labor supply, then the allocation of  $\{\ell, x\}$  is exactly the same as in the baseline model and does not depend on the value of  $\rho$ . Independently of preference, the qualitative path of the optimal labor wedge (tax) is the same as in the baseline model, as described in Figures 1 and 2.

Next, we consider the case with borrowing constraints on households (and the planner), so that  $\iota_b = 1$ , and the optimal path of  $\nu$  satisfies:

$$\dot{\nu} - \delta\nu = -\left[1 + \gamma(\nu - 1)\right]\frac{\alpha}{\eta}\frac{y}{x} + \left(\overbrace{(\varrho - \rho) - \frac{\dot{\mu}}{\mu}}^{=\psi > 0}\right)\nu,$$

where the last term on the right was previously absent. Given that the economy is growing and the planner is impatient,  $b \ge 0$  is binding, and consumption is output determined,  $c = (1 - \alpha)y(x, \ell)$ , and increases over time with wealth x accumulation. This, in turn, implies that  $u_c = \mu$  falls over time, and  $\psi = (\varrho - \rho) - \dot{\mu}/\mu > 0$  for any value of  $\varrho \ge \rho$ . The more impatient is the planner, the more binding is the constraint, and the larger is  $\psi$ . The presence of  $\psi > 0$  is equivalent to a larger discount rate  $\delta$ , making the accumulation of wealth x (and its contribution to future productivity) less valuable to the planner. This is a general effect from borrowing constraints on households, which is present independently of the present bias of the planner, however it gets amplified by the present bias  $\varrho - \rho > 0$ . Lastly, one can show that the long-run labor tax ( $\bar{\tau}_{\ell} > 0$ ) increases in  $\varrho$  relative to its baseline level, which is still optimal when  $\varrho = \rho$ , even under borrowing constraints. In all cases, it is still true that for low enough  $x_0, \nu(0) > 1$ , and the planner starts the transition with a labor subsidy, as in the baseline model. See illustration in Figure A8.



Figure A8: Households with finite lives and borrowing constraints Note: As in Figure 3.

## A4 Extensions

### A4.1 Persistent productivity types

Suppose that there are two types of entrepreneurs, H and L, and the analysis extends naturally to any finite number of types. Each type of entrepreneur draws its productivity from a Pareto distribution  $G_j(z) = 1 - (z/b_j)^{-\eta_j}$ , where  $b_j$  is a lower bound and  $\eta_j$  is the shape parameter, for  $j \in \{H, L\}$ , such that

$$\frac{\eta_H}{\eta_H - 1} b_H > \frac{\eta_L}{\eta_L - 1} b_H,$$

so that H-type entrepreneurs are more productive on average. The *j*-type entrepreneurs redraw their productivities *iid* from  $G_j(z)$  each instant, and at a certain rate they transition to another type over time. Specifically, entrepreneurs of type L switch to type H at a Poisson rate p and, conversely, type H entrepreneurs switch to type L at rate q (i.e., over any interval of time, the type distribution follows a Markov process).

Note that this way of modeling the productivity process maintains the tractability of our framework due to a continuous productivity distribution within types, yet allows us to accommodate an arbitrary amount of persistence in the productivity process over time. Indeed, by varying  $b_j$ ,  $\eta_j p$ and q, we can parameterize an arbitrary productivity process in terms of persistence: for example, with p = q = 0 and  $\eta_H = \eta_L \to \infty$ , we obtain perfectly persistent productivity types  $b_H > b_L$ .<sup>10</sup> In the rest of the analysis, we impose for simplicity  $\eta_H = \eta_L = \eta \in (1, \infty)$ .

Note that under this formulation, upon the realization of instantaneous productivity z, the within-period behavior of entrepreneurs is characterized by Lemma 1 independently of the type of the entrepreneur (i.e., independently of whether z was drawn from the H or the L distribution). Furthermore, the aggregation results in Lemma 2 still apply but within each productivity type, so that we can write in particular:

$$y_j = \Theta_j x_j^{\gamma} \ell_j^{1-\gamma}, \quad \text{where} \quad \Theta_j \equiv \frac{r^*}{\alpha} \left[ \frac{\lambda \eta b_j}{\eta - 1} \left( \frac{\alpha A}{r^*} \right)^{\eta/\alpha} \right]^{\gamma},$$

and  $y_j$ ,  $x_j$ ,  $\ell_j$  are the aggregate output, wealth and labor demand of entrepreneurs of type  $j \in \{L, H\}$ . The wealth dynamics now satisfy:

$$\dot{x}_{L} = \frac{\alpha}{\eta} y_{L}(x_{L}, \ell_{L}) + (r^{*} - \delta) x_{L} + q x_{H} - p x_{L},$$
(A33)

$$\dot{x}_{H} = \frac{\alpha}{\eta} y_{H}(x_{H}, \ell_{H}) + (r^{*} - \delta) x_{H} + q x_{L} - p x_{H},$$
(A34)

and the labor market clearing requires  $\ell_L + \ell_H = \ell$ , where  $\ell$  is labor supply in the economy.

To stay consistent with the spirit of our analysis, we consider the case in which the planner cannot tax differentially the L and H types of entrepreneurs, and in particular imposes a common labor income tax on the households, independently of which type of entrepreneur they are working for. Therefore, the additional constraint on the planner's implementation is the equalization of the

<sup>&</sup>lt;sup>10</sup>Nonetheless, we need to impose certain regularity conditions if we want to make use of the type of characterization as in Lemma 2, since we need to ensure that the least productive draws within each type remain inactive along the transition path.

marginal products of labor (and hence wages) across the two types of entrepreneurs:

$$\frac{(1-\alpha)y_L(x_L,\ell_L)}{\ell_L} = \frac{(1-\alpha)y_H(x_H,\ell_H)}{\ell_H} = w.$$
 (A35)

The household budget constraint can then be written as:

$$c + b = w(\ell_L + \ell_H) + r^*b.$$
 (A36)

Following the same steps as in Lemma 3, we can show that the planner maximizes household utility (1) (where  $\ell = \ell_L + \ell_H$ ) by choosing  $\{c, \ell_L, \ell_H, b, x_L, x_H, w\}$ , which satisfy (A33)–(A36), with the associated vector of Lagrange multipliers  $\mu \cdot (\nu_L, \nu_H, \xi_L, \xi_H, 1)'$ . Forming a Hamiltonian and taking the optimality conditions, we arrive after simplification at similar results as in (27)–(29), in particular (27) still holds, and we have:

$$-\frac{u_{\ell}}{u_c} = (1 - \tau_{\ell}) \frac{(1 - \alpha)y}{\ell}, \quad \text{where} \quad \tau_{\ell} \equiv \gamma (1 - \bar{\nu}), \tag{A37}$$

and  $y = y_L(x_L, \ell_L) + y_H(x_H, \ell_H), \ \ell = \ell_L + \ell_H$ , and

$$\bar{\nu} \equiv \frac{\ell_L \nu_L + \ell_H \nu_H}{\ell},$$

i.e.  $\bar{\nu}$  is an employment-weighted average of the co-states  $(\nu_L, \nu_H)$  for the state variables  $(x_L, x_H)$ .<sup>11</sup>

Lastly, we have two optimality conditions for  $x_j$ , which determine the dynamics of  $\nu_j$ , in parallel with (29):

$$\dot{\nu}_{L} = (\delta + p)\nu_{L} - q\nu_{H} - \left[\gamma\nu_{L} + (1 - \gamma)\xi_{L}\right]\frac{\alpha}{\eta}\frac{y_{L}(x_{L}, \ell_{L})}{x_{L}},\\ \dot{\nu}_{H} = (\delta + q)\nu_{H} - p\nu_{L} - \left[\gamma\nu_{H} + (1 - \gamma)\xi_{H}\right]\frac{\alpha}{\eta}\frac{y_{H}(x_{H}, \ell_{H})}{x_{H}}.$$

When  $(x_L, x_H)$  are both low, then  $(\nu_L, \nu_H)$  are both high, and so is  $\bar{\nu}$ , which means that the planner subsidizes labor. Therefore, our main results generalize immediately to the case with a persistent productivity process.

## A4.2 Pareto weight on entrepreneurs

Consider an extension to the planning problem (P1) in Section 3.2 in which the planner puts a positive Pareto weight  $\theta > 0$  on the utilitarian welfare criterion of all entrepreneurs  $\mathbb{V}_0 \equiv \int V_0(a) dG_{a,0}(a)$ ,

$$\begin{aligned} c: & u_c = \mu, \\ w: & \mu \ell = \mu \xi_L \ell_L + \mu \xi_H \ell_H, \\ \ell_j: & -u_\ell = \mu \frac{(1-\alpha)y_j}{\ell_j} \left[ 1 + \gamma(\nu_j - \xi_j) \right], \qquad j \in \{L, H\}, \end{aligned}$$

and we can sum the last condition for the two j's weighting by  $\ell_j$ , and manipulate using the other two conditions and (A35) to arrive at (A37).

<sup>&</sup>lt;sup>11</sup>The underlying FOCs are

where  $V_0(\cdot)$  is the expected value to an entrepreneur with initial assets  $a_0$ . From Appendix A1.3, we have:

$$\mathbb{V}_0 = v_0 + \frac{1}{\delta} \int \log a \, \mathrm{d}G_{a,0}(a) + \frac{1}{\delta} \int_0^\infty e^{-\delta t} \frac{\alpha}{\eta} \frac{y(x,\ell)}{x} \mathrm{d}t.$$

Since given the instruments the planner cannot affect the first two terms in  $\mathbb{V}_0$ , the planner's problem in this case can be written as:

$$\max_{\{c,\ell,b,x\}_{t\geq 0}} \int_0^\infty e^{-\rho t} u(c,\ell) dt + \frac{\theta}{\delta} \int_0^\infty e^{-\delta t} \frac{\alpha}{\eta} \frac{y(x,\ell)}{x} dt$$
(P7)  
subject to 
$$c + \dot{b} = (1-\alpha)y(x,\ell) + r^*b,$$
$$\dot{x} = \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x.$$

The Hamiltonian for this problem is:

$$\mathcal{H} = u(c,\ell) + \frac{\theta}{\delta} e^{-(\delta-\rho)t} \frac{\alpha}{\eta} \frac{y}{x} + \mu \left[ (1-\alpha)y(x,\ell) + r^*b - c \right] + \mu \nu \left[ \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x \right],$$

and the optimality conditions are  $u_c(t) = \mu(t) = \overline{\mu}$  for all t and:

$$\frac{\partial \mathcal{H}}{\partial \ell} = u_{\ell} + \bar{\mu} \left[ \frac{\theta}{\delta \bar{\mu}} e^{-(\delta - \rho)t} \frac{\gamma}{x} + (1 - \gamma) + \gamma \nu \right] (1 - \alpha) \frac{y}{\ell} = 0,$$
$$\dot{\nu} - \rho \nu = -\frac{1}{\bar{\mu}} \frac{\partial \mathcal{H}}{\partial x} = (\delta - r^*) \nu - \left[ \frac{\theta}{\delta \bar{\mu}} e^{-(\delta - \rho)t} \frac{\gamma}{x} + (1 - \gamma) + \gamma \nu \right] \frac{\alpha}{\eta} \frac{y}{x}.$$

The dynamic system characterizing  $(x, \nu)$  is the same as in Section 3.2 with the exception of an additional term  $\frac{\theta}{\delta \bar{\mu}} e^{-(\delta - \rho)t} \frac{\gamma}{x} \ge 0$  in the conditions above. Similarly, the optimal labor wedge which we denote by  $\tau_{\ell}^{\theta}$  is given by (32). Note that the long-run optimal tax rate is the same for all  $\theta \ge 0$ , as a consequence of our assumption that entrepreneurs are more impatient than workers,  $\delta > \rho$ . When  $\delta = \rho$ , the long-run tax depends on  $\theta$  and can be negative for  $\theta$  large enough.

## A4.3 Closed economy

We can also extend our analysis to the case of a closed economy in which the total supply of capital equals the sum of assets held by workers and entrepreneurs,  $\kappa(t) = x(t) + b(t)$ , and the interest rate, r(t), is determined endogenously to equalize the demand and supply of capital.<sup>12</sup> In what follows, we set up formally the closed economy model. In particular, we generalize Lemmas 2 and 3 to show that the constraints on allocations (25)–(26) in the closed economy become:

$$\dot{b} = \left[ (1 - \alpha) + \alpha \frac{\eta - 1}{\eta} \frac{b}{\kappa} \right] y(\kappa, x, \ell) - c - \varsigma_x x, \tag{A38}$$

$$\dot{x} = \left[1 + (\eta - 1)\frac{x}{\kappa}\right]\frac{\alpha}{\eta}y(\kappa, x, \ell) + (\varsigma_x - \delta)x,\tag{A39}$$

<sup>&</sup>lt;sup>12</sup>Another interesting case, which we do not consider here, is that of a large open economy, in which the optimal unilateral policy additionally factors in the incentives to manipulate the country's intra- and intertemporal terms of trade (see, for example, Costinot, Lorenzoni, and Werning, 2014).

and where the output function is now:

$$y(\kappa, x, \ell) = \Theta^c \left(\kappa^{\eta - 1} x\right)^{\alpha} \ell^{1 - \alpha} \quad \text{with} \quad \Theta^c \equiv A \left(\frac{\eta}{\eta - 1} \lambda^{1/\eta}\right)^{\alpha}, \tag{A40}$$

instead of (16). For generality, we allowed for transfers  $\varsigma_x$  between households and entrepreneurs, as in Appendix A2.2. The only other difference between (A38)–(A39) and (25)–(26) is that we have substituted in the expression for the equilibrium interest rate from (18),  $r = \alpha(\eta - 1)/\eta \cdot y/\kappa$ , which continues to hold in the closed economy. The closed economy dynamics depend on an additional state variable—the capital stock,  $\kappa$ .

We solve the planner's problem and characterize the optimal policies in the closed economy below. The main new result is that the planner no longer keeps the intertemporal margin undistorted, and chooses to encourage workers' savings in the early phase of transition, provided  $x/\kappa$  is low enough. This allows the economy to accumulate capital,  $\kappa$ , faster, which in turn raises output and profits, and speeds up entrepreneurial wealth accumulation. The long-run intertemporal wedge may be positive, negative or zero, depending on how large  $x/\kappa$  is in the steady state. The qualitative predictions for the labor wedge remain the same as in the small open economy: an initial labor supply subsidy is replaced eventually by a labor supply tax after entrepreneurs have accumulated enough wealth.

Lemma 1, as well as aggregation equations (13)–(15) and income accounting equations (18) from Lemma 2, still apply in the closed economy. The difference however is that now r is endogenous and we have an additional equilibrium condition  $\kappa = x + b$ . Substituting capital demand (13) into the aggregate production function (15), we obtain (A40) which defines  $y(x, \kappa, \ell)$  in the text. We can then summarize the planner's problem in the closed economy (without transfers,  $\varsigma_x \equiv 0$ ) as:

$$\max_{\{c,\ell,\kappa,b,x\}} \int_0^\infty e^{-\rho t} u(c,\ell) dt,$$
(PC)  
subject to  $\dot{b} = \left[ (1-\alpha) + \alpha \frac{\eta - 1}{\eta} \frac{b}{\kappa} \right] y(x,\kappa,\ell) - c,$   
 $\dot{x} = \left[ 1 + (\eta - 1) \frac{x}{\kappa} \right] \frac{\alpha}{\eta} y(x,\kappa,\ell) - \delta x,$ 

with  $\kappa_0 = x_0 + b_0$  given. Note that we have used (18) to substitute out the endogenous interest rate r.

To simplify notation, we replace the first constraint with the sum of the two constraints to substitute  $\dot{\kappa}$  for  $\dot{b} + \dot{x}$ , with  $\tilde{\mu}$  now denoting a co-state for  $\kappa$ . The Hamiltonian for this problem is:

$$\mathcal{H} = u(c,\ell) + \tilde{\mu} \left[ y(\kappa,\ell,x) - c - \delta x \right] + \tilde{\mu} \tilde{\nu} \left( \left[ 1 + (\eta - 1) \frac{x}{\kappa} \right] \frac{\alpha}{\eta} y(\kappa,\ell,x) - \delta x \right),$$

where  $\tilde{\nu}$  corresponds to  $\nu - 1$  in the baseline planner's problem (P1), as we have used the sum of the two budget constraint (country aggregate resource constraint) instead of using the household budget constraint ( $\dot{\kappa}$  vs.  $\dot{b}$ ). The optimality conditions are:

$$0 = \frac{\partial \mathcal{H}}{\partial c} = u_c - \tilde{\mu},$$
  
$$0 = \frac{\partial \mathcal{H}}{\partial \ell} = u_\ell + \tilde{\mu} \left[ 1 + \tilde{\nu} \frac{\alpha}{\eta} \left( 1 + (\eta - 1) \frac{x}{\kappa} \right) \right] (1 - \alpha) \frac{y}{\ell}$$

and

$$\dot{\tilde{\mu}} - \rho \tilde{\mu} = -\frac{\partial \mathcal{H}}{\partial \kappa} = -\tilde{\mu}r - \tilde{\mu}\tilde{\nu}\frac{\alpha}{\eta}\left(1 + (\eta - 1)\frac{x}{\kappa}\right)r + \tilde{\mu}\tilde{\nu}\frac{x}{\kappa}r,$$
$$(\dot{\tilde{\mu}\tilde{\nu}}) - \rho \tilde{\mu}\tilde{\nu} = -\frac{\partial \mathcal{H}}{\partial x} = -\tilde{\mu}\left(\frac{\alpha}{\eta}\frac{y}{x} - \delta\right) - \tilde{\mu}\tilde{\nu}\left(\frac{\alpha}{\eta}\left(1 + (\eta - 1)\frac{x}{\kappa}\right)\frac{\alpha}{\eta}\frac{y}{x} + r - \delta\right).$$

From the second condition we have the labor wedge:

$$-\frac{u_{\ell}}{u_{c}} = \left[1 + \tilde{\nu}\frac{\alpha}{\eta}\left(1 + (\eta - 1)\frac{x}{\kappa}\right)\right](1 - \alpha)\frac{y}{\ell} \qquad \Rightarrow \qquad \tau_{\ell}^{c} = -\tilde{\nu}\frac{\alpha}{\eta}\left(1 + (\eta - 1)\frac{x}{\kappa}\right).$$

Next we use the other conditions to characterize the intertemporal wedge:

$$\frac{\dot{u}_c}{u_c} = \rho - r - \tilde{\nu}r \left[ \frac{\alpha}{\eta} - \frac{x}{\kappa} \left( 1 - \alpha \frac{\eta - 1}{\eta} \right) \right] \qquad \Rightarrow \qquad \tau_b^c = -\tilde{\nu}r \left[ \frac{\alpha}{\eta} - \frac{x}{\kappa} \left( 1 - \alpha \frac{\eta - 1}{\eta} \right) \right]$$

Finally, we have:

$$\dot{\tilde{\nu}} = \left(\delta + \tilde{\nu}r\left[\frac{\alpha}{\eta} - \frac{x}{\kappa}\left(1 - \alpha\frac{\eta - 1}{\eta}\right)\right] - \frac{\alpha}{\eta}\left(1 + (\eta - 1)\frac{x}{\kappa}\right)\frac{\alpha}{\eta}\frac{y}{x}\right)\tilde{\nu} - \left(\frac{\alpha}{\eta}\frac{y}{x} - \delta\right).$$

This dynamic system can be solved using conventional methods to show that the optimal tax on labor supply and worker savings satisfy:

$$\tau_{\ell}^{c}(t) = -\left(1 + (\eta - 1)\frac{x(t)}{\kappa(t)}\right)\frac{\alpha}{\eta}\tilde{\nu}(t) \quad \text{and} \quad \tau_{b}^{c}(t) = -r(t)\left(1 - \frac{1}{\gamma}\frac{x(t)}{\kappa(t)}\right)\frac{\alpha}{\eta}\tilde{\nu}(t),$$

where  $\tilde{\nu}(t) = \nu(t) - 1$  is again a co-state for x(t).

## A4.4 Optimal intertemporal wedge

Assume the planner cannot manipulate the labor supply margin, and can only distort the intertemporal margin. The planner's problem in this case can be written as:

$$\max_{\{c,\ell,b,x\}_{t\geq 0}} \int_0^\infty e^{-\rho t} u(c,\ell) dt$$
(P6)  
subject to 
$$c + \dot{b} = (1-\alpha)y(x,\ell) + r^*b,$$
$$\dot{x} = \frac{\alpha}{\eta}y(x,\ell) + (r^* - \delta)x,$$
$$-\frac{u_c}{u_\ell} = (1-\alpha)\frac{y(x,\ell)}{\ell},$$

where the last constraint implies that the planner cannot distort labor supply, and we denote by  $\mu\psi$  the Lagrange multiplier on this additional constraint. We can write the Hamiltonian for this problem as:

$$\mathcal{H} = u(c,\ell) + \mu \left[ (1-\alpha)y(x,\ell) + r^*b - c \right] + \mu \nu \left[ \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x \right] + \mu \psi \left[ (1-\alpha)y(x,\ell) - h(c,\ell) \right],$$

where  $h(c, \ell) \equiv -\ell u_{\ell}(c, \ell)/u_{c}(c, \ell)$ . The optimality conditions are:

$$0 = \frac{\partial \mathcal{H}}{\partial c} = u_c - \mu (1 + \psi h_c),$$
  

$$0 = \frac{\partial \mathcal{H}}{\partial \ell} = u_\ell + \mu (1 - \gamma + \gamma \nu) (1 - \alpha) \frac{y}{\ell} + \mu \psi \left( (1 - \gamma) (1 - \alpha) \frac{y}{\ell} - h_\ell \right),$$
  

$$\dot{\mu} - \rho \mu = -\frac{\partial \mathcal{H}}{\partial b} = -\mu r^*,$$
  

$$(\dot{\mu}\nu) - \rho \mu \nu = -\frac{\partial \mathcal{H}}{\partial x} = -\mu (1 - \gamma + \gamma \nu) \frac{\alpha}{\eta} \frac{y}{x} - \mu \nu (r^* - \delta) - \mu \psi (1 - \gamma) \frac{\alpha}{\eta} \frac{y}{x}.$$

Under our parameter restriction  $\rho = r^*$ , the third condition implies  $\dot{\mu} = 0$  and  $\mu(t) \equiv \bar{\mu}$  for all t, however, now  $u_c = \bar{\mu}(1 + \psi h_c)$  and is no longer constant in general, reflecting the use of the savings subsidy to workers. Combining this with the second optimality condition and the third constraint on the planner's problem, we have:

$$(1-\alpha)\frac{y}{\ell} = -\frac{u_c}{u_\ell} = \frac{((1-\gamma)(1+\psi) + \gamma\nu)(1-\alpha)y/\ell - \psi h_\ell}{1+\psi h_c}$$

which we simplify using  $h = (1 - \alpha)y$ :

$$\psi = \frac{\gamma(\nu - 1)}{h_c + \ell h_\ell / h - (1 - \gamma)}.$$
(A41)

Finally, the dynamics of  $\nu$  satisfy:

$$\dot{\nu} = \delta\nu - \left((1-\gamma)(1+\psi) + \gamma\nu\right)\frac{\alpha}{\eta}\frac{y}{x},$$

and the distortion to the consumption smoothing satisfies:

$$u_c = \bar{\mu}(1 + \psi h_c) = \bar{\mu}(1 + \Gamma(\nu - 1)), \qquad \Gamma \equiv \frac{\gamma h_c}{h_c + \ell h_\ell / h - (1 - \gamma)}.$$
 (A42)

Recall that under  $\rho = r^*$ ,  $\dot{u}_c/u_c = -\varsigma_b$ , and therefore  $\varsigma_b > 0$  whenever  $\psi h_c = \Gamma(\nu - 1)$  is decreasing over time.

## A5 Appendix for Quantitative Model in Section 4

This appendix describes the quantitative model of Section 4 in more detail. It lays out the entire system of equations that constitute an equilibrium with taxes. For simplicity, we first focus on the case with only a labor tax  $\tau_{\ell}(t)$  and abstract from other tax instruments. The generalization to other tax instruments is straightforward, and we detail the specific case of a credit subsidy below.

**Workers** The first-order condition of workers is the same as in the baseline model, namely (24). Their budget constraint is given by:

$$c(t) = (1 - \tau_{\ell}(t))w(t)\ell(t) + T(t).$$
(A43)
**Entrepreneurs** As explained in the main text, it now becomes necessary to keep track of the joint distribution of entrepreneurial wealth a and productivity z. To this end, denote by g(a, z, t) the density corresponding to the CDF  $\mathcal{G}_t(a, z)$ . The problem of an entrepreneur still separates into a static profit maximization and a dynamic consumption-saving problem. Profits are given by:

$$\pi(a, z; w(t), r^*) = \max_{n \ge 0, k \le \lambda a} \left\{ z(k^{\alpha} n^{1-\alpha})^{\beta} - w(t)n - r^*k \right\}.$$
 (A44)

As explained in Achdou, Han, Lasry, Lions, and Moll (2017), any dynamic optimization problem with a continuum of agents (like the one here) can be formulated and solved in terms of a system of two PDEs: a Hamilton-Jacobi-Bellman equation and a Kolmogorov Forward equation.

To keep the notation manageable, denote by  $\mu(z) = \left(-\nu \log z + \frac{\sigma^2}{2}\right) z$  and  $\sigma^2(z) = \sigma^2 z^2$  the drift and diffusion coefficients of the process for z corresponding to (37).<sup>13</sup> With this notation in hand, the system of PDEs summarizing entrepreneurs' behavior is:

$$\rho v(a, z, t) = \max_{c} u(c) + \partial_{a} v(a, z, t) [\pi(a, z; w(t), r^{*}) + r^{*}a - c] + \partial_{z} v(a, z, t) \mu(z) + \frac{1}{2} \partial_{zz} v(a, z, t) \sigma^{2}(z) + \phi \int_{\underline{z}}^{\overline{z}} (v(a, x, t) - v(a, z, t)) p(x) dx + \partial_{t} v(a, z, t), \partial_{t} g(a, z, t) = - \partial_{a} [s(a, z, t)g(a, z, t)] - \partial_{z} [\mu(z)g(a, z, t)] + \frac{1}{2} \partial_{zz} [\sigma^{2}(z)g(a, z, t)]$$
(A46)

$$-\phi g(a, z, t) + \phi p(z) \int_{z} g(a, x, t) dx,$$
  

$$s(a, z, t) = \pi(a, z; w(t), r^{*}) + r^{*}a - c(a, z, t),$$
(A47)

with initial condition  $g(a, z, 0) = g_0(a, z)$  and terminal condition  $\lim_{T\to\infty} v(a, z, T) = v_{\infty}(a, z)$ where  $v_{\infty}$  is the solution to the stationary analogue of (A45)–(A47). The value function satisfies the boundary conditions corresponding to reflecting barriers at  $\underline{z}$  and  $\overline{z}$ , namely  $\partial_z v(a, \underline{z}, t) =$  $\partial_z v(a, \overline{z}, t) = 0$  for all (a, t). For numerical reasons, we also impose a state constraint  $a \ge 0$ and therefore impose the corresponding state-constraint boundary condition (see Achdou, Han, Lasry, Lions, and Moll, 2017). The last two terms in the Kolmogorov Forward equation (A46) capture inflows and outflows due to Poisson productivity shocks: at rate  $\phi$ , individuals switch to another productivity types and hence the outflow term  $-\phi g(a, z, t)$ ; conversely, individuals with other productivity types (of which there are a mass  $\int_{\underline{z}}^{\overline{z}} g(a, x, t) dx$ ) switch to productivity type zat rate  $\phi p(z)$  and hence the inflow term  $+\phi p(z) \int_{\underline{z}}^{\overline{z}} g(a, x, t) dx$ .<sup>14</sup>

<sup>&</sup>lt;sup>13</sup>Formally, the infinitesimal generators  $\mathcal{A}f(z) := \lim_{t\downarrow 0} \frac{\mathbb{E}[f(z_t)] - f(z)}{t}$  of the productivity process (37) is given by  $\mathcal{A}f(z) = \mu(z)f'(z) + \frac{1}{2}\sigma^2(z)f''(z) + \phi \int_{\underline{z}}^{\overline{z}} (f(x) - f(z))p(x)dx$ , with  $\mu(z)$  and  $\sigma^2(z)$  defined in the text, and where p is the density of a truncated Pareto distribution with tail parameter  $\eta$ . The boundary conditions are those corresponding to reflecting barriers:  $f'(z) = f'(\overline{z}) = 0$ . The expression for the drift  $\mu(z)$  follows from Ito's formula: if  $\log z_t$  follows (37), then the drift of  $z_t$  must be  $(-\nu \log z_t + \sigma^2/2)z_t$ .

<sup>&</sup>lt;sup>14</sup>It is also straightforward to show using integration by parts that the operator in the Kolmogorov Forward equation  $(\mathcal{A}^*g)(z) = -\phi g(z) + \phi p(z) \int_{\underline{z}}^{\overline{z}} g(x) dx$  is the adjoint of the operator in the HJB equation ("infinitesimal generator")  $(\mathcal{A}f)(z) = \phi \int_{\underline{z}}^{\overline{z}} (f(x) - f(z)) p(x) dx$ , i.e.  $\langle \mathcal{A}f, g \rangle = \langle f, \mathcal{A}^*g \rangle$  where  $\langle \cdot, \cdot \rangle$  denotes the inner product, e.g.  $\langle f, g \rangle = \int_{\overline{z}}^{\overline{z}} f(x)g(x) dx$ .

**Government** As explained in the main text, we restrict the tax function to the parametric functional form (38). The government further runs a balanced budget and hence

$$T(t) = \tau_{\ell}(t)w(t)\ell(t). \tag{A48}$$

**Equilibrium** The equilibrium wage w(t) clears the labor market:

$$\omega \int_{\underline{z}}^{\overline{z}} \int_{0}^{\infty} n(a, z; w(t), r^{*}) g(a, z, t) dadz = (1 - \omega)\ell(t).$$
(A49)

Given initial condition  $g_0(a, z)$ , the two PDEs (A45), (A46) together with (A47), workers' optimality conditions (24) and (A43), the government budget constraint (A48) and the equilibrium condition (A49) fully characterize equilibrium.

**Optimal Policy** The optimal tax policy is found as follows. For any triple of parameters  $(\underline{\tau}_{\ell}, \bar{\tau}_{\ell}, \gamma_{\ell})$ , we can compute a time-dependent equilibrium by solving the system of equations laid out above. Given this, we compute welfare  $\mathcal{V}_0(\underline{\tau}_{\ell}, \bar{\tau}_{\ell}, \gamma_{\ell})$  defined in (39) and we find the triple  $(\underline{\tau}_{\ell}, \bar{\tau}_{\ell}, \gamma_{\ell})$  that maximizes this objective function. To do this in practice, we simply discretize the three tax parameters using discrete grids and employ a simple grid search to find the optimum.

Welfare Measure To measure the welfare gain of switching from the laissez-faire equilibrium to optimal policy (or any other policy), we use a standard consumption-equivalent welfare metric which we denote by  $\Delta$ . Denoting the equilibrium allocation under laissez-faire with hats,  $\Delta$  solves

$$(1-\omega)\int_0^\infty e^{-\rho t} u((1+\Delta)\widehat{c}_t,\widehat{\ell}_t) dt + \omega\theta \int \mathbb{E}_0 \left[\int_0^\infty e^{-\delta t} \log((1+\Delta)\widehat{c}_t^e) dt \Big| (a_0,z_0) = (a,z)\right] d\widehat{\mathcal{G}}_0(a,z) = (1-\omega)\int_0^\infty e^{-\rho t} u(c_t,\ell_t) dt + \omega\theta \int \mathbb{E}_0 \left[\int_0^\infty e^{-\delta t} \log(c_t^e) dt \Big| (a_0,z_0) = (a,z)\right] d\mathcal{G}_0(a,z) = \mathcal{V}_0,$$

where we use that the t = 0 value to entrepreneurs is  $v_0(a, z) = \mathbb{E}_0 \left[ \int_0^\infty e^{-\delta t} \log(c_t^e) dt \Big| (a_0, z_0) = (a, z) \right]$ . Since  $u((1 + \Delta)c, \ell) = \log(1 + \Delta) + \log c - \ell^{1+\varphi}/(1 + \varphi)$ , the last equation can be written as

 $\left(\frac{1-\theta}{\rho} + \frac{\theta}{\delta}\right)\log(1+\Delta) + \widehat{\mathcal{V}}_0 = \mathcal{V}_0$  and hence

$$\Delta = \exp(\bar{\rho}(\mathcal{V}_0 - \widehat{\mathcal{V}}_0)) - 1, \quad \text{where} \quad \bar{\rho} \equiv \left(\frac{1-\theta}{\rho} + \frac{\theta}{\delta}\right)^{-1}.$$

This is the number reported in the first column of Table 1 (with different rows corresponding to different policy experiments). Similarly, workers' and entrepreneurs' consumption-equivalent welfare changes are given by

$$\Delta^w = \exp(\rho(\mathcal{V}_0^w - \widehat{\mathcal{V}}_0^w)) - 1, \qquad \Delta^e = \exp(\delta(\mathcal{V}_0^e - \widehat{\mathcal{V}}_0^e)) - 1,$$

where  $\mathcal{V}_0^w = \int_0^\infty e^{-\rho t} u(c_t, \ell_t) dt$  and  $\mathcal{V}_0^e = \int_{\underline{z}}^{\overline{z}} \int_0^\infty v_0(a, z) g_0(a, z) dadz$  and similarly for  $\widehat{\mathcal{V}}_0^w$  and  $\widehat{\mathcal{V}}_0^e$ . These numbers for workers are reported in the second column of Table 1, and the numbers for entrepreneurs are omitted for brevity. **Parameterization** Table A2 reports the parameter values we use in our quantitative exercise. A number of these were already discussed in the main text. We here provide additional detail and discuss the values of those parameters not discussed in the main text.

Param.	Value	Description	Comment/Source
$\chi$	0.1	scaling factor of initial dist.	initial $=1/10$ stationary wealth
$\lambda$	2	tightness of financial constraint	steady state $D/Y = 2.29$
$\alpha$	0.33	capital share	standard value
$\beta$	0.90	returns to scale	De Loecker et al. $(2016)$
$e^{-\nu}$	0.85	autocorrelation of productivity	Asker et al. $(2014)$
$\sigma$	0.3	innovation variance of productivity	Asker et al. $(2014)$
$\phi$	0.1	arrival rate of Poisson shocks	jump on average every 10 yrs
$\zeta$	1.1	Pareto tail of Poisson shocks	
$\bar{z}/\underline{z}$	7.33	upper/lower productivity bound	
ρ	0.03	discount rate of workers	set equal to $r^*$
$\delta$	0.05	discount rate of entrepreneurs	
$1/\varphi$	1	Frisch elasticity	Blundell et al. $(2016)$
ω	1/3	population share of entrepreneurs	typical developing country
$r^*$	0.03	world interest rate	standard value

Table A2: Parameter Values for Quantitative Exercise

As stated in the main text, we set the initial wealth-productivity distribution  $\mathcal{G}_0(a, z)$  equal to the stationary distribution in the absence of policy  $\mathcal{G}_{\infty}(a, z)$  but with every entrepreneur's wealth scaled down by a factor of ten. More precisely, we parameterize the initial distribution as  $\mathcal{G}_0(a, z) = \mathcal{G}_{\infty}(a/\chi, z), 0 < \chi < 1$  for all a so that, in particular, aggregate initial wealth is a fraction  $\chi$  of aggregate final wealth  $\int a d\mathcal{G}_0(a, z) = \int a d\mathcal{G}_{\infty}(a/\chi, z) = \chi \int a d\mathcal{G}_{\infty}(a, z)$ . We then set  $\chi = 0.1$ . The capital share  $\alpha$  and returns to scale  $\beta$  are set to standard parameter values from the literature (e.g. Atkeson and Kehoe, 2007). In fact, an estimate for returns to scale of 0.9 is on the lower end of the estimates of De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) for India (see their Table 3). There are two reasons for choosing such a low value of the returns-to-scale parameter  $\beta$ . First, our aim is to be conservative and to show the robustness of our results to sizable deviations from constant returns to scale. Second, even though empirical estimates of production functions typically find values of  $\beta$  close to one, firms may face downward-sloping demand curves, thereby resulting in a revenue function that has lower returns to scale than the (physical) production function and so  $\beta = 0.9$  may not be unreasonable.

As discussed in the text, the parameters of the productivity process are calibrated following the estimates of Asker, Collard-Wexler, and de Loecker (2014).<sup>15</sup> Also as discussed in the text,  $\lambda$  is calibrated to match the ratio of external finance to GDP. We define external finance as the sum of private credit, private bond market capitalization, and stock market capitalization in the data of Beck, Demirguc-Kunt, and Levine (2000). This definition follows Buera, Kaboski, and Shin (2011; see also their footnote 9). In our model, the external finance to GDP ratio is given by  $D_t/Y_t$ , where  $D_t = \int \max\{k_t(a, z) - a, 0\} d\mathcal{G}_t(a, z)$  and  $Y_t = \int y_t(a, z) d\mathcal{G}_t(a, z)$ .

<sup>&</sup>lt;sup>15</sup>See their Online Appendix Table OA.11 at http://www.princeton.edu/~jdeloeck/ACWDLappendix.pdf.

As in the baseline model, we set workers' discount rate  $\rho$  equal to the world interest rate which we set to  $r^* = 0.03$ . Again as in the baseline model, a stationary distribution only exists if the entrepreneurial discount rate  $\delta$  exceeds the interest rate  $r^*$ . We set  $\delta = 0.05$  resulting in a gap between workers' and entrepreneurs' discount rates of  $\delta - \rho = 0.02$ . We set the Frisch elasticity governing workers' labor supply decision to one. This number is slightly higher than the 0.82 identified by Chetty, Guren, Manoli, and Weber (2011) as the representative estimate from existing studies of the micro elasticity at the individual level, accounting for intensive and extensive margins of adjustment. At the household level though, the marginally attached worker is often the wife (at least in developed countries) and a Frisch labor supply elasticity of one is in line with the estimates of Blundell, Pistaferri, and Saporta-Eksten (2016) for married women.

**Myopic Labor Union** Consider a labor union that restricts labor supply  $\ell$  to maximize current worker utility  $u(c, \ell)$ , where  $c = w\ell$  given that workers are borrowing constrained, and the union internalizes its effect on the equilibrium wage rate w. The optimality condition for the union is:

$$u_c \cdot \left[ w + \frac{\partial w}{\partial \ell} \ell \right] + u_\ell = 0 \qquad \Rightarrow \qquad -\frac{u_\ell}{u_c} = \left[ \underbrace{1 - 1/\varepsilon_{LS}}^{\equiv 1 - \hat{\tau}_\ell^U} \right] w,$$

where  $\varepsilon_{LS} \equiv -\frac{\partial \ell}{\partial w} \frac{w}{\ell}$  is the aggregate labor demand elasticity by the entrepreneurs. In particular, we have the aggregate labor demand given by:

$$\ell = \ell_t(w; r^*) = \int n(a, z; w, r^*) \mathrm{d}\mathcal{G}_t(a, z),$$

where  $n(a, z; w, r^*)$  is the labor demand policy function of individual entrepreneurs which maximizes profit (A44) and it satisfies:

$$n(a, z; w, r^*) = \left[\frac{(1-\alpha)\beta A \left(zk(a, z; w, r^*)\right)^{\alpha\beta}}{w}\right]^{\frac{1}{1-(1-\alpha)\beta}},$$
  
where  $k(a, z; w, r^*) = \min\left\{\lambda a, \left[\left(\frac{\alpha}{r^*}\right)^{1-(1-\alpha)\beta} \left(\frac{1-\alpha}{w}\right)^{(1-\alpha)\beta} \beta A z^{\alpha\beta}\right]^{\frac{1}{1-\beta}}\right\}.$ 

Under these circumstances, we can calculate the aggregate labor demand elasticity:

$$\varepsilon_{LS} = \frac{1}{1 - (1 - \alpha)\beta} + \frac{\alpha\beta}{1 - \beta} \frac{(1 - \alpha)\beta}{1 - (1 - \alpha)\beta} \pi_{+}$$

where  $\pi \in [0, 1]$  is the share of labor hired by unconstrained entrepreneurs:

$$\pi = \pi_t(w; r^*) \equiv \frac{1}{\ell_t(w; r^*)} \int_{k(a, z; w, r^*) > \lambda a} n(a, z; w, r^*) \mathrm{d}\mathcal{G}_t(a, z).$$

To summarize, the myopic union tax  $\hat{\tau}_{\ell}^{U} = 1/\varepsilon_{ES}$  is decreasing in the elasticity of the aggregate labor demand (just like a monopoly markup), which in turn is increasing in the fraction of unconstrained entrepreneurs (who are more elastic because they can adjust capital).

Numerically, we solve for the time path of the union tax as a dynamic fixed point jointly with



Figure A9: Optimal policy in quantitative model: credit subsidy  $\varsigma_k(t)$  and GDP Y(t)Note: In panel (b), steady-state GDP in the laissez-faire equilibrium is normalized to 1.

the wage rate and the labor share of the unconstrained entrepreneurs,  $\{\hat{\tau}^U_{\ell}(t), w(t), \pi(t)\}_{t\geq 0}$ . For a given path of  $\{\hat{\tau}^U_{\ell}(t), w(t)\}$  we solve for equilibrium dynamics and recover the path of  $\{\pi(t)\}$ , which we use to update the union tax schedule, and iterate until convergence.

**Credit Subsidy** Denote the credit subsidy by  $\varsigma_k(t)$ . Analogously to (38) we assume that  $\varsigma_k(t)$  is a parametric function of time:

$$\varsigma_k(t) = e^{-\gamma_k t} \cdot \underline{\varsigma}_k + \left(1 - e^{-\gamma_k t}\right) \cdot \overline{\varsigma}_k.$$

The economy is the same as above except for three changes. First, we obviously set  $\tau_{\ell}(t) = 0$  for all t. Second, we replace (A44) by

$$\pi(a, z; w(t), r^*, \varsigma_k(t)) = \max_{n \ge 0, k \le \lambda a} \left\{ z(k^{\alpha} n^{1-\alpha})^{\beta} - w(t)n - (1 - \varsigma_k(t))r^*k \right\}.$$

Third, we replace the government budget constraint (A48) by

$$(1-\omega)T(t) + \omega\varsigma_k(t)r^* \int k_t(a,z) \mathrm{d}\mathcal{G}_t(a,z) = 0.$$

Analogously to above, we search for a triple of parameters  $(\underline{\varsigma}_k, \overline{\varsigma}_k, \gamma_k)$  that maximizes welfare  $\mathcal{V}_0$  defined in (39). Figure A9 reports the results in an analogous fashion to Figure 4. The resulting welfare effects for the case  $\theta = 1/2$  are reported in Table 1 in the main text.

# A6 Analysis of the Multi-sector Model in Section 5

# A6.1 Setup of the multi-sector economy

We start our analysis with three types of tax instruments: a savings tax, sector-specific consumption taxes, and sector-specific labor income taxes. These taxes directly distort the actions of the households, while they have only an indirect effect on the entrepreneurs through market prices, namely sector-specific wage rates and output prices. We briefly discuss below the extension of our analysis to production, credit and export subsidies.

**Households** The households have general preferences over n + 1 goods,  $u = u(c_0, c_1, \ldots, c_n)$ , where good i = 0 is traded internationally, and we choose it as numeraire, normalizing  $p_0 = 1$ . Goods  $i = 1, \ldots, n$  may be tradable or non-tradable, and we assume for concreteness that goods  $i = 1, \ldots, k$  are tradable and goods  $i = k + 1, \ldots, n$  are not tradable. The households maximize the intertemporal utility given by  $\int_0^\infty e^{-\rho t} u(t) dt$ , and supply inelastically a total of L units of labor, which is split between the sectors:<sup>16</sup>

$$\sum_{i=0}^{n} \ell_i = L. \tag{A50}$$

The after-tax wage across all sectors must be equalized in order for the households to supply labor to every sector:

$$(1 - \tau_i^{\ell})w_i = w, \qquad i = 0, 1, \dots n,$$
 (A51)

where w is the common after-tax wage,  $w_i$  is the wage paid by the firms in sector i and  $\tau_i^{\ell}$  is the tax on labor income earned in sector i.<sup>17</sup>

The households have access to a risk-free instantaneous bond which pays out in the units of the numeraire good i = 0, and face the following budget constraint:

$$\sum_{i=0}^{n} (1+\tau_i^c) p_i c_i + \dot{b} \le (r-\tau^b) b + wL + T,$$
(A52)

where  $p_i$  is the producer price of and  $\tau_i^c$  is the consumption tax on good *i*, *b* is the asset position of the households,  $\tau^b$  is a savings tax, and *T* is the lump-sum transfer from the government. The solution to the household problem is given by the following optimality conditions:

$$\begin{cases} \frac{\dot{u}_0}{u_0} = \rho + \tau^b + \frac{\dot{\tau}_0^c}{1 + \tau_0^c} - r, \\ \frac{u_i}{u_0} = \frac{1 + \tau_i^c}{1 + \tau_0^c} p_i, \quad i = 1, \dots, n, \end{cases}$$
(A53)

where  $u_i \equiv \partial u / \partial c_i$  is the marginal utility from consumption of good *i*. The first condition is the Euler equation for the intertemporal allocation of consumption. The second set of conditions is the optimal intratemporal consumption choice across sectors. It is easy to see that one of the taxes is redundant, and we normalize  $\tau_0^c \equiv 0$  in what follows.

<sup>&</sup>lt;sup>16</sup>Note that the assumption of inelastic labor supply is without loss of generality since we can always choose sector i = n to be non-tradable (so that  $c_n = y_n$ ) with competitive production according to  $y_n = \ell_n$ . This is equivalent to either home production or leisure, generalizing the setup of Section 2 of the paper.

<sup>&</sup>lt;sup>17</sup>One of the labor taxes is redundant, and we can normalize  $\tau_0^{\ell} \equiv 0$  (or alternatively  $\tau_n^{\ell} = 0$ , if the labor allocated to this sector is interpreted as leisure), but we find it more convenient to keep this extra degree of freedom in characterizing the optimal wedges.

**Production** The production in each sector is carried out by heterogeneous entrepreneurs, as in Section 2. Entrepreneurs within each sector face sector-specific collateral constraints parameterized by  $\lambda_i$ , operate sector-specific Cobb-Douglas technologies with productivity  $A_i$  and capitalintensity  $\alpha_i$ , and draw their idiosyncratic productivities from sector-specific Pareto distributions with tail parameter  $\eta_i$ . As a result, and analogous to Lemma 2, the aggregate production function in sector *i* is:<sup>18</sup>

$$y_i(x_i, \ell_i; p_i) = p_i^{\gamma_i(\eta_i - 1)} \Theta_i x_i^{\gamma_i} \ell_i^{1 - \gamma_i}, \qquad (A54)$$

where

$$\gamma_i = \frac{\alpha_i/\eta_i}{1 - \alpha_i + \alpha_i/\eta_i}$$
 and  $\Theta_i = \frac{r}{\alpha_i} \left[ \frac{\eta_i \lambda_i}{\eta_i - 1} \left( \frac{\alpha_i A_i}{r} \right)^{\eta_i/\alpha_i} \right]^{\gamma_i}$ 

Note that the producer price enters the reduced-form output function. A higher sectoral price allows a greater number of entrepreneurs to profitably produce, affecting both the production cutoff  $\underline{z}_i$  and the amount of capital  $\kappa_i$  used in the sector, which enter the sectoral production function corresponding to (15).

Sectoral entrepreneurial wealth  $x_i$  is in the units of the numeraire good i = 0, which we think of as the capital good in the economy. Further, we assume that entrepreneurs consume only the capital good i = 0, so that with logarithmic utility aggregate consumption of sector i entrepreneurs is  $\delta x_i$ , where  $\delta$  is the entrepreneurial discount rate.<sup>19</sup> As a result, the evolution of sectoral entrepreneurial wealth satisfies:

$$\dot{x}_i = \frac{\alpha_i}{\eta_i} p_i y_i(x_i, \ell_i; p_i) + (r - \delta) x_i, \tag{A55}$$

where as before  $\alpha_i/\eta_i$  is the share of profits in the sectoral revenues, and  $(1 - \alpha_i)$  is the share of labor income:

$$w_i \ell_i = (1 - \alpha_i) p_i y_i(x_i, \ell_i; p_i).$$
(A56)

**Government** The government chooses the tax policy  $(\tau^b, \{\tau_i^c, \tau_i^\ell\}, T)_{t \ge 0}$  and runs a balanced budget:

$$T = \tau^b b + \sum_{i=0}^n \left( \tau_i^c p_i c_i + \tau_i^\ell w_i \ell_i \right).$$
(A57)

Note that we rule out direct sectoral transfers which would allow the planner to effectively sidestep the financial constraints.<sup>20</sup>

**Prices** We consider here a small open economy which takes the price of capital  $r^* = \rho$  as given, as well as the international prices of the tradable goods  $(p_0, p_1, \ldots, p_k)$  for  $k \leq n$ . The prices of the non-tradables  $p_i$  for  $i = k + 1, \ldots, n$  are determined to clear the respective markets:

$$c_i = y_i(x_i, \ell_i; p_i), \qquad i = k + 1, \dots, n.$$
 (A58)

<sup>&</sup>lt;sup>18</sup>Production function (A54) and income accounting (A56) follow from the same derivation as Lemma A5 in Appendix A3.3, since price  $p_i$  plays an equivalent role to output subsidy  $\varsigma_y$  in that derivation.

<sup>&</sup>lt;sup>19</sup>We make this assumption for tractability, but the analysis extends to more general utility functions of entrepreneurs.

<sup>&</sup>lt;sup>20</sup>In the presence of unbounded transfers, the planner instantaneously jumps every sector to its optimal steady-state level of financial wealth  $\bar{x}_i^*$ , while no other policy instrument is used, just as in the one-sector economy in Appendix A2.2.

### A6.2 Optimal policy

We now analyze optimal policy in this framework. With the structure above, we can prove the following *primal approach* lemma, which generalizes the earlier Lemma 3 (and its proof follows the same steps):

**Lemma A6** Given initial condition b(0) and  $\{x_i(0)\}$ , for any allocation  $(\{c_i, \ell_i, x_i\}_{i=0}^N, b)_{t\geq 0}$  that satisfies the following dynamic system:

$$\begin{cases} \dot{b} = r^* b + \sum_{i=0}^{N} \left[ (1 - \alpha_i) p_i y_i(x_i, \ell_i; p_i) - p_i c_i \right], \\ \dot{x}_i = \frac{\alpha_i}{\eta_i} p_i y_i(x_i, \ell_i; p_i) + (r^* - \delta) x_i, \quad i = 0, 1, \dots, N, \\ c_i = y_i(x_i, \ell_i; p_i), \quad i = k+1, \dots, N, \\ L = \sum_{i=0}^{N} \ell_i, \end{cases}$$
(A59)

there exists a path of taxes  $\{\tau_i^b, \{\tau_i^c, \tau_i^\ell\}_{i=0}^n, T\}_{t\geq 0}$  that decentralize this allocation as an equilibrium in the multi-sector economy, where  $r^*$  and  $(p_0, \ldots, p_k)$  are international prices and  $(p_{k+1}, \ldots, p_n)$  can be chosen by the planner along with the rest of the allocation.

Therefore, we can consider a planner who maximizes household utility with respect to

$$\left\{b, \{c_i, \ell_i, x_i\}_{i=0}^n, \{p_i\}_{i=k+1}^n\right\}_{t>0}$$

and subject to the set of constraint in (A59) with corresponding Lagrange multipliers denoted by  $\mu \cdot (1, \{\nu_i\}_{i=0}^n, \{\psi_i\}_{i=k+1}^n, \omega)$ . Given a dynamic allocation, we recover the corresponding paths of taxes  $\tau^b$  and  $\{\tau_i^c, \tau_i^\ell\}$  from household optimality (A53) together with sectoral labor demand which satisfies (A56). Lastly, note that the availability of consumption taxes allows the planner to create a wedge between the sectoral production and consumption prices, and in the non-tradable sectors this allows the planner to manipulate equilibrium producer prices.

We now outline and discuss some general results, and in the following sections consider two illustrative special cases to explore in more detail the implications for comparative advantage and the real exchange rate. After considering the problem with the full set of instruments, we consecutively limit the set of taxes available to the planner. First, we additionally rule out the sectoral labor taxes ( $\tau_i^{\ell} \equiv 0$  for all *i*) by imposing

$$(1 - \alpha_i)p_i y(x_i, \ell_i; p_i) = w\ell_i, \qquad i = 0, 1, \dots, n,$$
 (A60)

in addition to constraints in (A59), and we denote the corresponding Lagrange multipliers by  $\mu \xi_i$ for  $i = 0, \ldots, n.^{21}$  Second, we also rule out static consumption taxes ( $\tau_i^c \equiv 0$  for all *i*), leaving the planner with only the intertemporal tax  $\tau^b$ . For this case, we hence additionally impose

$$u_i = p_i u_0, \qquad i = 1, \dots, n,$$
 (A61)

and denote by  $\mu \chi_i$  the corresponding Lagrange multipliers.

We prove the following result which applies to both the case with the full set of instruments and the cases with limited instruments (see Appendix A6.5 below for formal derivation):

<sup>&</sup>lt;sup>21</sup>Note that in this case the common wage rate w becomes a variable of planner's optimization.

**Lemma A7** (a) The planner never uses consumption taxes on tradable goods ( $\tau_i^c \equiv 0$  for  $i = 0, \ldots, k$ ); (b) The planner does not use the intertemporal tax ( $\tau^b \equiv 0$ ) as long as static sectoral taxes (labor and/or consumption) are available.

The planner never uses consumption taxes on tradable goods because they only distort consumption and have no effect on producers, who face unchanged international prices.<sup>22</sup> As in the analysis of the one-sector economy of Section 3.2, the planner does not distort the intertemporal margin (the Euler equation of households) as long as she has access to some static sectoral instruments, either consumption or labor income taxes. Indeed, such instruments are more direct, operating immediately over the sectoral allocation of resources, which is affected only indirectly by the intertemporal allocation of consumption through income effects. This implies that the widespread policies of 'financial repression' and government reserve accumulation can only be third-best in a small open economy with financial frictions, and would only be used in the absence of static sectoral instruments (as we discuss in more detail in Appendix A6.4, along with the implications for the real exchange rate).

Next consider the case where the planner has at her disposal both sectoral labor and consumption taxes. We show:

**Proposition A6** The optimal consumption and labor taxes in the multi-sector economy are given by:

$$\tau_i^c = \begin{cases} 0, & i = 0, 1, \dots, k, \\ \frac{1}{\eta_i - 1} (1 - \nu_i), & i = k + 1, \dots, n, \end{cases} \quad and \quad \tau_i^\ell = \begin{cases} \gamma_i (1 - \nu_i), & i = 0, 1, \dots, k, \\ -\tau_i^c, & i = k + 1, \dots, n, \end{cases}$$

where  $\nu_i$  is the shadow value of entrepreneurial wealth in sector *i*.

The planner does not tax consumption of tradables (as was already pointed out in Lemma A7), but does tax the consumption of non-tradables in proportion with  $(1 - \nu_i)$ . In other words, the planner subsidizes the consumption of non-tradables in sectors that have  $\nu_i > 1$ , meaning that they are financially constrained, and this subsidy is larger the more fat-tailed is the distribution of sectoral productivities (the smaller is  $\eta_i$ ). When tradable sectors are financially constrained,  $\nu_i > 1$ , the planner instead subsidizes labor supply to these sectors,  $\tau_i^{\ell} < 0$ , generalizing the result in a one-sector economy in (30). In contrast, the labor tax for the non-tradable sectors perfectly offsets the corresponding consumption subsidies,  $\tau_i^{\ell} = -\tau_i^c$  for i > k.

To understand the overall effect of these various tax instruments, it is useful to define the *overall* labor wedge for sector i as:

$$1 + \tau_i \equiv \frac{(1 - \alpha_i)\frac{u_i y_i}{\ell_i}}{(1 - \alpha_0)\frac{u_0 y_0}{\ell_0}} = (1 - \tau_0^{\ell})\frac{1 + \tau_i^c}{1 - \tau_i^{\ell}}.$$

In words, the overall labor wedge is the combination of the product-market wedge  $1 + \tau_i^c$ , capturing deviations of consumers' marginal rate of substitution from relative sectoral prices, and the labor-market wedge  $(1-\tau_0^\ell)/(1-\tau_i^\ell)$  capturing deviations of the economy's marginal rate of transformation from relative prices. When the overall labor wedge is positive, the planner diverts the allocation of

 $<sup>^{22}</sup>$ The situation is different if the planner has access to production or export taxes for tradable goods, which we discuss below.

labor away from sector i (relative to the numeraire sector), and vice versa. Using Proposition A6, for tradable sectors we have

$$\tau_i = \frac{\tau_i^{\ell} - \tau_0^{\ell}}{1 - \tau_i^{\ell}} = \frac{\gamma_0(\nu_0 - 1) - \gamma_i(\nu_i - 1)}{1 + \gamma_i(\nu_i - 1)}, \qquad i = 1, \dots, k$$
(A62)

and for non-tradable sectors

$$\tau_i = -\tau_0^{\ell} = \gamma_0(\nu_0 - 1), \qquad i = k + 1, \dots, N.$$
(A63)

Consider first the overall labor wedge for non-tradables in (A63). Somewhat surprisingly, it is shaped exclusively by the need to subsidize the *tradable* sectors (in particular, the numeraire sector, which we chose as the base, since the wedges are relative by definition). This is because the need for financing in the non-tradable sector is addressed with respective consumption taxes,  $\tau_i^c$ . In other words, the presence of both consumption and production taxes for non-tradable sectors allows the planner to subsidize the non-tradable producers via an increase in producer prices  $p_i$  (due to  $\tau_i^c < 0$ ) without distorting the labor supply to these sectors. This option is unavailable in the tradable sectors which face exogenous international producer prices. Consider next the overall labor wedge for tradables in (A62). The allocation of labor is distorted in favor of the tradable sector *i* (relative to numeraire sector 0), that is  $\tau_i < 0$ , whenever  $\gamma_i(\nu_i - 1) > \gamma_0(\nu_0 - 1)$ , and vice versa. In the following Appendixes A6.3 and A6.4 we consider special cases in which we can further characterize the conditions under which certain sectors are subsidized or taxed.

The results here generalize to the case with a larger set of policy instruments. Specifically, when credit and/or output (export) subsidies are available, the planner optimally combines them with the labor subsidies to the constrained sectors according to the values of  $\nu_i$ . The planner wants to use all of these instruments in tandem to achieve the best outcome with minimal distortions, as we showed in Section 3.3 and Appendix A3.3 in the context of a one-sector economy. The advantage of output (export) subsidies over consumption subsidies in the tradable sectors is that they directly change effective producer prices even when the international prices are taken as given.

### A6.3 Comparative advantage and industrial policies

Proposition A6 characterizes policy in a general multi-sector economy in terms of planner's shadow values  $\nu_i$ , which represent the tightness of sectoral financial constraints. To make further progress in characterizing the policy in terms of the primitives of the economy, we consider in turn a few illuminating special cases. In this subsection we focus on the economy with tradable sectors only. For simplicity, we focus on two tradable sectors i = 0, 1, but the results extend straightforwardly to an economy with any number  $k \geq 2$  of tradable sectors.

First, we consider the case in which sectors are symmetric in everything except in what we call their *latent*, or long-run, comparative advantage. In particular, we assume that  $\eta_i \equiv \eta$  and  $\alpha_i \equiv \alpha$ for both sectors, and as a result  $\gamma_i \equiv \gamma$ . In this case, from (A54), sectoral revenues which also determine wages and profits are given by  $p_i y_i = p_i^{\zeta} \Theta_i x_i^{\gamma} \ell_i^{1-\gamma}$  where  $\zeta \equiv 1 + \gamma(\eta - 1)$ . We define a sector's latent comparative advantage to be the effective revenue productivity term  $p_i^{\zeta} \Theta_i$ . As reflected in its definition,  $\Theta_i$  may differ across sectors due to either physical productivity  $A_i$  or financial constraints  $\lambda_i$ , which for example depend on the pledgeability of sectoral assets. Importantly, a sector's *actual*, or short-run, comparative advantage may differ from this latent comparative advantage: in particular, it is also shaped by the allocation of sectoral entrepreneurial wealth  $x_i$  and is given by  $p_i^{\zeta} \Theta_i x_i^{\gamma}$ . In the short run, the country may specialize against its latent comparative advantage, if entrepreneurs in that sector are poorly capitalized (as was pointed out in Wynne, 2005). In the long run, the latent comparative advantage forces dominate, and entrepreneurial wealth relocates towards the sector with the highest  $p_i^{\zeta} \Theta_i$ .

We can apply the results of Proposition A6 to this case. In particular, using (A62) we have:

$$\tau_1 = \frac{\gamma(\nu_0 - \nu_1)}{1 + \gamma(\nu_1 - 1)},$$

and the planner shifts labor towards sector 0 whenever  $\nu_0 > \nu_1$ . We prove below that a necessary and sufficient condition for this is that sector 0 possesses a long-run comparative advantage, i.e.  $p_0^{\zeta}\Theta_0 > p_1^{\zeta}\Theta_1$ , independently of the initial allocation of wealth  $x_0$  and  $x_1$ , and hence short-run export patterns. We illustrate the optimal policy and resulting equilibrium dynamics relative to laissezfaire in Figure 5 in the text of the paper. The planner distorts the market allocation, and instead of equalizing marginal revenue products of labor across the two sectors, tilts the labor supply towards the latent comparative advantage sector. This is because the planner's allocation is not only shaped by the current labor productivity, which is increasing in wealth  $x_i$ , but also takes into account the shadow value of the sectoral entrepreneurial wealth, which depends on the latent comparative advantage  $p_i^{\zeta}\Theta_i$ . To summarize, the planner favors the long-run comparative advantage sector and speeds up the reallocation of factors towards it, consistent with some popular policy prescriptions (see e.g. Lin, 2012, and other references in the Introduction).

Second, we briefly consider the case in which sectors are asymmetric in terms of their structural parameters  $\alpha_i$  and  $\eta_i$ . To focus attention on this asymmetry, we shut down the comparative advantage forces just analyzed, so that a laissez-faire steady state features diversification of production across sectors.<sup>23</sup> We prove in Appendix A6.5 that the planner in this case nonetheless chooses to "pick a winner" by subsidizing one of the two sectors and independently of the initial conditions drives the economy to long-run specialization in this sector. Furthermore, there also exist cases in which the laissez-faire economy specializes in one sector, but the planner chooses to reverse the pattern of specialization.<sup>24</sup>

<sup>23</sup>The wage rate paid by sector i in the long run equals (see Appendix A6.5):

$$w = \left(\frac{\gamma_i}{1 - \gamma_i}\right)^{\frac{\gamma_i}{1 - \gamma_i}} \frac{\left[(1 - \alpha_i)p_i^{\zeta_i}\Theta_i\right]^{\frac{1}{1 - \gamma_i}}}{(\delta - \rho)^{\frac{\gamma_i}{1 - \gamma_i}}}$$

When the parameter combination on the right-hand side of this expression is equalized across sectors i = 0, 1, no sector has comparative advantage in the long run. That is, there exists a multiplicity of steady states without specialization, and the specific steady state reached (in terms of intersectoral allocation of labor) depends on the initial conditions. In the alternative case, the economy specializes in the long run in the sector for which this parameter combination is largest.

<sup>24</sup>In the long-run, somewhat counterintuitively, the planner drives the economy toward specialization in the sector with the *lower*  $\gamma_i$ . The intuition for this result can be obtained from the one-sector economy in Section 3, and in particular the formula for the steady-state tax (31). As explained there, the planner taxes rather than subsidizes entrepreneurs in steady state. As can be seen from (31), the size of this tax is increasing in  $\gamma$ . This is because a higher  $\gamma$  implies a larger "monopoly tax effect", i.e. a higher desire to redistribute from entrepreneurs to workers. This intuition carries over to the multi-sector economy studied here, and the planner puts a higher steady-state tax on the sector with higher  $\gamma_i$ , thereby specializing against it in the long run. Things may be different during the transition.

#### A6.4 Non-tradables, the real exchange rate and competitiveness

We now analyze in more detail a second case with only two sectors: a tradable sector i = 0and a non-tradable sector i = 1. This special case allows us to characterize more sharply the optimal sectoral taxes and particularly the implications for the real exchange rate. We find it useful to distinguish between two different measures of the real exchange rate: first, the CPIbased real exchange rate which in our two-sector model is pinned down by the after-tax price of non-tradables,  $(1 + \tau_1^c)p_1$ ; and, second, the wage-based real exchange rate which can be viewed as a measure of the country's competitiveness.<sup>25</sup> We will show below that optimal policies have potentially different implications for the two measures of real exchange rates. In particular, what happens to the CPI-based real exchange rate depends on the instruments at the planner's disposal.

All tax instruments We first consider the case where the planner has at her disposal the whole set of tax instruments we started with in Section 5. In this case, Proposition A6 applies and from (A63) the overall labor wedge is given by:

$$\tau_1 = \gamma_0(\nu_0 - 1),$$

which is positive whenever the tradable sector is undercapitalized, that is  $\nu_0 > 1$ . Hence labor is diverted away from non-tradables to tradables and, since production features decreasing returns to labor, wages paid by tradable producers  $w_0 = (1 - \alpha_0)y_0/\ell_0$  are compressed. The implications for the after-tax price of non-tradables and hence the CPI-based real exchange rate are more subtle. Since the consumer price of non-tradables is  $p_1(1 + \tau_1^c) = u_1/u_0$ , one needs to understand the behavior of marginal utilities relative to the competitive equilibrium. A complete characterization is possible in the limit case when capital intensity in the non-tradable sector becomes very small  $\alpha_1 \rightarrow 0$ , and hence non-tradable production is frictionless. We prove in Appendix A6.5 that in this case, the CPI-based real exchange rate necessarily *appreciates*. Intuitively, labor is reallocated towards tradables and hence non-tradable production decreases. Since non-tradables become more scarce, their price increases and hence the CPI-based real exchange rate appreciates. In numerical experiments (omitted for brevity), we have computed time paths for the equilibrium allocation in the case with  $\alpha_1 > 0$ , which indicate that also in this case the CPI-based real exchange rate is *appreciated* relative to the competitive equilibrium when the tradable sector is sufficiently undercapitalized.

No sectoral labor taxes Sector-specific labor taxes might be unavailable to the planner if it is hard to allocate jobs and occupations to specific sectors in order to administer such taxes. We thus consider the case where the planner cannot differentially tax labor in different sectors. Since labor supply is inelastic in our multi-sector economy, this means that the planner cannot directly affect the allocation of labor at all. Therefore, the only instrument used by the planner is the consumption tax in the non-tradable sector,  $\tau_1^c$ , since according to Lemma A7 neither the savings subsidy, nor the tradable consumption tax are used. Indeed, we prove in Appendix A6.5 that the

<sup>&</sup>lt;sup>25</sup>The CPI-based real exchange rate is given by  $P/P^*$ , where P and  $P^*$  are the price indexes of the home country and the rest of the world which are functions of the consumer prices of tradable and non-tradable goods. Since we analyze a small open economy,  $P^*$  is fixed from the point of view of the home country, and we normalize  $p_0 = 1$  and  $\tau_0^c = 0$ . Therefore, the real exchange rate appreciates whenever the consumer price of non-tradables  $(1 + \tau_1^c)p_1$  increases. The wage-based real exchange rate is given by  $w/w^*$ , where  $w^*$  is the wage rate in the rest of the world and is taken as given.

planner only uses the non-tradable consumption tax and sets it according to:

$$\tau_1^c = \frac{1}{\eta_1/\alpha_1 - 1} \left[ (1 - \nu_1) + \frac{1 - \gamma_1}{\gamma_1} \kappa \right], \quad \text{where} \quad \kappa = \frac{(\nu_0 - 1)\ell_0 - \frac{1}{\eta_1/\alpha_1 - 1}(\nu_1 - 1)\ell_1}{\ell_0 + \frac{\eta_1 - 1}{\eta_1/\alpha_1 - 1}\ell_1}.$$

The expression for the non-tradable tax depends on two terms. The first term is similar to above and the planner subsidizes non-tradables whenever the sector is undercapitalized, i.e.  $\nu_1 > 1$ . In contrast, the second term  $\kappa$  captures the fact that the planner uses the consumption tax to also affect the labor allocation. Note that  $\kappa$  increases in  $\nu_0$  and decreases in  $\nu_1$ . Intuitively, if  $\nu_0$  is large, then the only way to improve the allocation is by taxing non-tradable consumption (which is reflected in the  $\kappa$  term in  $\tau_1^c$ ), thereby shifting labor to the tradable sector. From the expression above we see that non-tradable consumption is taxed (i.e.,  $\tau_1^c > 0$ ) when:

$$\frac{\ell_0}{L}(\nu_0 - 1) > \frac{\gamma_1}{1 - \gamma_1}(\nu_1 - 1),$$

which is more likely the larger is  $\nu_0$ , the smaller is  $\nu_1$ , the larger is the size of the tradable sector 0 (in terms of labor allocated to this sector), and the smaller is  $\gamma_1$ . In particular, as  $\gamma_1 \rightarrow 0$  (for example, due to  $\alpha_1 \rightarrow 0$ , i.e. as non-tradable production stops relying on capital), non-tradable consumption is taxed whenever  $\nu_0 > 1$ . As a result, the economy-wide wage  $w = (1 - \alpha)y_0/\ell_0$  decreases and hence the wage-based real exchange rate  $w/w^*$  depreciates. At the same time, non-tradables become more expensive due to the consumption tax, and hence the *CPI*-based real exchange rate appreciates.

No sectoral taxes In the absence of any sectoral instruments (labor or consumption), the planner has to recur to intertemporal distortions by means of a savings subsidy, or a policy of capital controls and reserve accumulation more commonly used in practice (see Jeanne, 2013, for the equivalence result of these policies). We provide a formal analysis of this case below, and here briefly discuss the results. We show that by taxing consumption today in favor of future periods, the planner shifts resources away from the non-tradable sector and towards the tradable sector, which is desirable when  $\nu_0$  is sufficiently large. The effect of such policy on the allocation of labor across sectors is similar to that of a consumption tax on non-tradables. However, it comes with an additional intertemporal distortion on the consumption of tradables, and as a result the intertemporal policy is strictly dominated by static sectoral policies (as follows from Lemma A7). In response to the intertemporal policy, wages, the price of non-tradables, and consumption of both goods decrease, while the tradable sector expands production and exports, facing unchanged international prices.<sup>26</sup> Both CPI- and wage-based real exchange rates depreciate in response to this policy, which contrasts with the previously discussed cases. This narrative is consistent with the analysis of Song, Storesletten, and Zilibotti (2014) who argue that, in China, a combination of capital controls and other policies compressed wages and increased the wealth of entrepreneurs, thereby relaxing their borrowing constraints.

To summarize the analysis of this section, one of the goals of the planner is to shift labor towards the tradable sector when it is financially constrained ( $\nu_0 > 1$ ), which can be achieved in a variety of ways depending on the available set of instruments. One common feature of the policies is that they

<sup>&</sup>lt;sup>26</sup>Net exports result in net foreign asset accumulation, which is however accompanied by an inflow of productive capital to satisfy increased capital demand in the tradable sector.

reduce the equilibrium wage rate paid in the tradable sector, resulting in a depreciated wage-based real exchange rate and enhanced competitiveness of the tradable-sector firms. At the same time, the effect of the policies on the consumption prices and CPI-based real exchange rate depends on the available policy instruments. In particular, the planner favors static sectoral instruments, which tax non-tradable labor or consumption and result in appreciated non-tradable prices. The standard narrative in the optimal exchange rate policy literature (see Rodrik, 2008; Korinek and Servén, 2016; Benigno and Fornaro, 2012) focuses on the alternartive case where static sectoral taxes are unavailable, and the planner is limited to an intertemporal instrument. Thus, the real exchange rate implications of the optimal policy crucially depend on which instruments are available, even when the nature of inefficiency remains the same. We conclude that the (standard CPI-based) real exchange rate may not be a particularly useful guide for policymakers because there is no robust theoretical link between this variable and growth-promoting policy interventions.

# A6.5 Optimality conditions in the multi-sector economy

The planner's problem in the multi-sector economy can be summarized using the following Hamiltonian:

$$\begin{aligned} \mathcal{H} &= u(c_0, c_1, \dots, c_n) + \mu \left( rb + \sum_{i=1}^n \left[ (1 - \alpha_i) p_i y_i(x_i, \ell_i; p_i) - p_i c_i + \tau_i^x x_i \right] \right) \\ &+ \mu \sum_{i=1}^n \nu_i \left( \frac{\alpha_i}{\eta_i} p_i y_i(x_i, \ell_i; p_i) + (r - \delta - \tau_i^x) x_i \right) \\ &+ \mu \sum_{i=k+1}^n \psi_i (y_i(x_i, \ell_i; p_i) - c_i) + \mu \omega \left( L - \sum_{i=1}^n \ell_i \right) \\ &+ \mu \sum_{i=1}^n \xi_i \left( (1 - \alpha_i) p_i y_i(x_i, \ell_i; p_i) - \frac{w}{1 - \tau_i^\ell} \ell_i \right), \end{aligned}$$

with  $p_0 \equiv 1$ ,  $i \leq k$  tradable goods ( $p_i$  given exogenously) and i > k non-tradable goods ( $p_i$  determined in equilibrium to clear the goods market). The constraints with co-state variables  $\mu$  and  $\mu\nu_i$  correspond to the dynamic constraints on the evolution of the state variables b and  $x_i$ . The constraints with Lagrange multipliers  $\mu\psi_i$  and  $\mu\omega$  correspond to market clearing for non-tradable goods and for labor respectively. The last set of constraints with Lagrange multipliers  $\mu\xi_i$  additionally impose equalization of marginal products of labor across sectors, i.e. correspond to the case when sectoral labor taxes are ruled out ( $\tau_i^{\ell} \equiv 0$ ), and otherwise  $\xi_i \equiv 0$ . Finally, note that  $\tau_i^x$  are the sector-specific transfers of wealth from households to entrepreneurs.

We consider three special cases:

- 1. Most restrictive:  $\tau_i^x \equiv 0$  and  $\tau_i^\ell \equiv 0$
- 2. Baseline:  $\tau_i^x \equiv 0$ . When  $\tau_i^{\ell}$  are available and unconstrained, we have  $\xi_i \equiv 0$  (follows from the FOC for  $\tau_i^{\ell}$ ), and hence we simply drop the last line of constraints.
- 3. With transfers, i.e. with all instruments. Just like in the previous case, we drop the last line of constraints.

We consider cases in reverse order. There is also another case in which we rule out consumption taxes and impose  $u_i/u_0 = p_i$  for all *i*, which we consider separately below. In all three cases, the FOC for *b* implies  $\dot{\mu} = \mu(\rho - r) = 0$  and the FOC for  $c_0$  implies  $u_0 = \mu$ , hence the intertemporal

tax is not used,  $\tau^b \equiv 0.$  The FOCs for all other  $c_i{\rm 's}$  are

$$u_i = \mu(p_i + \psi_i),$$

where for convenience we introduce  $\psi_i \equiv 0$  for tradable  $i \leq k$ . We rewrite:

$$\frac{u_i}{u_0} = p_i + \psi_i = p_i(1 + \tau_i^c), \qquad \tau_i^c \equiv \frac{\psi_i}{p_i},$$

and  $\tau_i^c \equiv 0$  for  $i \leq k$ . The static FOCs for  $\ell_i$  for all i and  $p_i$  for i > k are:

$$\begin{pmatrix} (1-\alpha_i)(1+\xi_i) + \frac{\alpha_i}{\eta_i}\nu_i + \tau_i^c \end{pmatrix} (1-\gamma_i) \frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega + \xi_i \frac{w}{1-\tau_i^\ell}, \quad \forall i \\ \left( (1-\alpha_i)(1+\xi_i) + \frac{\alpha_i}{\eta_i}\nu_i \right) \left[ 1+\gamma_i(\eta_i-1) \right] + \gamma_i(\eta_i-1)\tau_i^c = 1, \quad i > k.$$

With transfers In this case, the FOC wrt  $\tau_i^x$  implies  $\nu_i \equiv 1$ . We also consider the case with  $\tau_i^\ell$  available, so that  $\xi_i \equiv 0$ . Therefore, we can rewrite the two FOCs above as:

$$\underbrace{\left(1 + \frac{1 - \gamma_i}{1 - \alpha_i}\tau_i^c\right)(1 - \alpha_i)\frac{p_iy_i(x_i, \ell_i; p_i)}{\ell_i} = \omega, \quad \forall i}_{=1 \text{ from definition of } \gamma_i} \underbrace{\left(1 - \alpha_i + \frac{\alpha_i}{\eta_i}\right)\left[1 + \gamma_i(\eta_i - 1)\right]}_{=1 \text{ from definition of } \gamma_i} + \gamma_i(\eta_i - 1)\tau_i^c = 1, \quad i > k.$$

Therefore, we have  $\tau_i^c \equiv 0$  for all i > k, and hence  $\tau_i^\ell \equiv 0$  for all i.

**Baseline** without transfers and with labor taxes. We still have  $\xi_i \equiv 0$ , but  $\nu_i \neq 1$  in general. Therefore, the two sets of conditions are:

$$\begin{pmatrix} 1 + \gamma_i(\nu_i - 1) + \frac{1 - \gamma_i}{1 - \alpha_i}\tau_i^c \end{pmatrix} (1 - \alpha_i)\frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega, \quad \forall i \\ 1 + \gamma_i(\nu_i - 1) + \gamma_i(\eta_i - 1)\tau_i^c = 1, \quad i > k, \end{cases}$$

where we use the property that:

$$1 + \gamma_i(\eta_i - 1) = \frac{1 - \gamma_i}{1 - \alpha_i}.$$

We simplify:

$$(1 + \gamma_i(\nu_i - 1)) (1 - \alpha_i) \frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega, \qquad i \le k,$$
  
$$(1 + \tau_i^c) (1 - \alpha_i) \frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega, \qquad i > k,$$
  
$$\tau_i^c = \frac{1 - \nu_i}{\eta_i - 1}, \qquad i > k.$$

We hence have:

$$\begin{aligned} \tau_i^\ell &= \gamma_i (1 - \nu_i), \qquad \tau_i^c \equiv 0, \qquad i \le k, \\ \tau_i^\ell &= -\tau_i^c, \qquad \tau_i^c = \frac{1 - \nu_i}{\eta_i - 1}, \qquad i > k, \end{aligned}$$

Lastly, we characterize the overall labor wedge in sector i:

$$1 + \tau_i = \frac{(1 - \alpha_i)\frac{u_i y_i(x_i, \ell_i; p_i)}{\ell_i}}{(1 - \alpha_0)\frac{u_0 y_0(x_0, \ell_0; p_0)}{\ell_0}} = \begin{bmatrix} \frac{1 + \gamma_0(\nu_0 - 1)}{1 + \gamma_i(\nu_i - 1)}, & i \le k, \\ 1 + \gamma_0(\nu_0 - 1), & i > k. \end{bmatrix}$$

This completes the proofs of Lemma A7 and Proposition A6 in Appendix A6.

**No labor taxes** In this case  $\tau_i^{\ell} \equiv 0$  and we have  $\xi_i \neq 0$ , and an additional FOC wrt w. We rewrite the full set of FOCs:

$$(1 - \gamma_i)(1 + \xi_i) + \gamma_i \nu_i + \frac{1 - \alpha_i}{1 - \gamma_i} \tau_i^c = 1 + \kappa + \xi_i, \quad \forall i (1 - \gamma_i)(1 + \xi_i) + \gamma_i \nu_i + \gamma_i (\eta_i - 1) \tau_i^c = 1, \quad i > k, \sum_{i=1}^n \xi_i \ell_i = 0,$$

where in the first set of FOCs we used that  $(1 - \alpha_i)p_iy_i/\ell_i = w$  for all i and replaced variables:

$$\kappa \equiv \frac{\omega - w}{w}.$$

Subtract the second line from the first line for i > k to get:

$$\begin{aligned} \tau_i^c &\equiv 0, \qquad i \le k, \\ \tau_i^c &= \kappa + \xi_i, \qquad i > k, \end{aligned}$$

where the first line simply restates the definition. Using these, we solve for  $\xi_i$ :

$$\begin{aligned} \xi_i &= (\nu_i - 1) - \frac{\kappa}{\gamma_i}, & i \le k, \\ \xi_i &= \frac{(1 - \nu_i) - (\eta_i - 1)\kappa}{\eta_i / \alpha_i - 1}, & i > k, \end{aligned}$$

so that

$$\tau_i^c = \frac{1}{\eta_i/\alpha_i - 1} \left[ (1 - \nu_i) + \frac{1 - \gamma_i}{\gamma_i} \kappa \right], \qquad i > k.$$

Therefore, the consumption tax on non-tradables decreases in  $\nu_i$  and increases in  $\kappa$ , which measures the average scarcity of capital across sectors. Derive the expression for  $\kappa$  from the last FOC for w:

$$\sum_{i \le k} \left[ (\nu_i - 1) - \frac{\kappa}{\gamma_i} \right] \ell_i + \sum_{i > k} \frac{(1 - \nu_i) - (\eta_i - 1)\kappa}{\eta_i / \alpha_i - 1} \ell_i = 0,$$

which implies:

$$\kappa = \frac{\sum_{i \le k} (\nu_i - 1)\ell_i - \sum_{i > k} \frac{1}{\eta_i / \alpha_i - 1} (\nu_i - 1)\ell_i}{\sum_{i \le k} \ell_i + \sum_{i > k} \frac{\eta_i - 1}{\eta_i / \alpha_i - 1} \ell_i}.$$

Therefore,  $\kappa$  increases in  $\nu_i$  for  $i \leq k$  and decreases in  $\nu_i$  for i > k. Specializing to the case with two sectors, a tradable sector 0 and a non-tradable sector 1, we obtain the results in Section A6.4.

We omit the discussion of the remaining case without static instruments ( $\tau_i^c = \tau_i^\ell = 0$ ) for brevity, and refer the reader to the analysis in Appendix A4.4 for the case with two sectors and  $\alpha_1 = 0$  in the non-tradable sector.

**Comparative advantage** We now specialize the analysis to the case with two tradable sectors with  $\alpha_i \equiv \alpha$  and  $\eta_i \equiv \eta$ , and hence  $\gamma_i = \gamma$ . We have  $\tau_i = \gamma(1 - \nu_i)$  for i = 0, 1. The system characterizing the planner's allocation is given by (for i = 0, 1):

$$\ell_0 + \ell_1 = L,$$

$$(1 + \tau_i)p_i^{1+\gamma(\eta-1)}\Theta_i \left(\frac{x_i}{\ell_i}\right)^{\gamma} = \omega,$$

$$\frac{\dot{x}_i}{x_i} = \frac{\alpha}{\eta}p_i^{1+\gamma(\eta-1)}\Theta_i \left(\frac{\ell_i}{x_i}\right)^{1-\gamma} + r - \delta,$$

$$\dot{\tau}_i = \delta + \frac{\delta}{\gamma}\tau_i - (1 + \tau_i)\frac{\alpha}{\eta}p_i^{1+\gamma(\eta-1)}\Theta_i \left(\frac{\ell_i}{x_i}\right)^{1-\gamma},$$

for some aggregate  $\omega$ , which is a function of time, and where  $\tau_i$  is a subsidy to labor in sector *i*. Solving for labor allocation:

$$\ell_i = \left[\frac{(1+\tau_i)p_i^{1+\gamma(\eta-1)}\Theta_i}{\omega}\right]^{1/\gamma} x_i.$$

and therefore labor balance implies:

$$\omega = \left(\frac{1}{L}\sum_{i} \left[(1+\tau_i)p_i^{1+\gamma(\eta-1)}\Theta_i\right]^{1/\gamma} x_i\right)^{\gamma}.$$

With this we are left with an autonomous system in  $\{x_i, \tau_i\}$ , which is almost separable across *i* and only connected by the aggregate variable  $\omega$ :

$$\frac{\dot{x}_i}{x_i} = \frac{\alpha}{\eta} \frac{(1+\tau_i)^{\frac{1-\gamma}{\gamma}} Z_i^{1/\gamma}}{W} + r - \delta,$$
$$\dot{\tau}_i = \delta + \frac{\delta}{\gamma} \tau_i - \frac{\alpha}{\eta} \frac{\left[(1+\tau_i) Z_i\right]^{1/\gamma}}{W},$$

where

$$Z_i \equiv p_i^{1+\gamma(\eta-1)}\Theta_i, \qquad W \equiv \omega^{\frac{1-\gamma}{\gamma}}.$$

In the laissez-faire allocation, the same system applies, but with  $\tau_i \equiv 0$ . The dynamics for this system are determined by  $Z_i$ , a measure of the latent comparative advantage. Both laissez-faire and planner's solution have unique and identical steady states with complete specialization, provided

that  $Z_0 \neq Z_1$ . If, for concreteness  $Z_0 > Z_1$ , then  $\bar{\ell}_1 = \bar{x}_1 = 0$  in the steady state,  $\bar{\ell}_0 = L$  and

$$\bar{x}_0 = \left[\frac{\alpha}{\eta} \frac{Z_0}{\delta - \rho}\right]^{\frac{1}{1 - \gamma}} L.$$

The planner sets a greater subsidy  $\tau_i$  for a sector with a greater latent comparative advantage  $Z_i$ , thus shifting labor allocation towards this sector and speeding up the transition towards the long-run equilibrium with specialization in this sector, as in Figure 5.

# A6.6 Overlapping production cohorts

Consider an immediate extension of our model (with firms paying out dividends when they die rather than with entrepreneurs) with a single homogeneous good but multiple cohorts producing it. Any living firm dies at a Poisson rate  $\delta$  and pays out its wealth back to households, while a new inflow of firms happens at rate  $\delta$  endowed with exogenous wealth  $\underline{x}_t$ . The productivity of new cohorts increases at rate g,  $\Theta_t = \Theta_0 e^{gt}$ . We write the production function of cohort  $s \in (-\infty, t]$  as:

$$y_s = y_s(x_s, \ell_s) = \Theta_s x_s^{\gamma} \ell_s^{1-\gamma}$$

Labor must be allocated between different cohorts splitting the exogenously given labor supply L:

$$\int_{-\infty}^{t} \ell_s \mathrm{d}s = L. \tag{A64}$$

In the decentralized equilibrium, labor is allocated according to:

$$\omega = (1 - \alpha) \frac{y_s}{\ell_s} = (1 - \alpha) \Theta_s \left(\frac{x_s}{\ell_s}\right)^{\gamma}, \tag{A65}$$

where  $\omega$  is the common wage rate. Therefore,  $\ell_s$  is allocated in proportion to short-run marginal product which is proportional to  $\Theta_s x_s^{\gamma}$ . The wealth of cohort s evolves according to:

$$\dot{x}_s = \frac{\alpha}{\eta} y_s + (r - \delta) x_s,$$

where  $\delta x_s$  is the transfer of net worth from exiting firms to households. The initial condition is  $x_t(t) = \underline{x}_t$ , and we parameterize it as a function of t in what follows. Lastly, the budget constraint of the households is given by:

$$\dot{b} = rb + \int_{-\infty}^{t} \left( (1-\alpha)y_s + \delta x_s \right) \mathrm{d}s - c - \underline{x}_t.$$

**Laissez-faire** There are only two dimensions of choice: 1) intertemporal allocation of consumption and 2) allocation of labor across cohorts  $\{\ell_s\}$ . The former one is undistorted, as in our baseline model. The latter solves (A64)–(A65), so that labor is allocated according to:

$$\ell_s = \frac{\Theta_s^{1/\gamma} x_s}{\int_{-\infty}^t \Theta_{\tilde{s}}^{1/\gamma} x_{\tilde{s}} \mathrm{d}\tilde{s}} L.$$

Therefore, the SR comparative advantage  $\Theta_s x_s^{\gamma}$  is a sufficient statistic for labor allocation.

**Aggregation** Given this decentralized labor allocation we can aggregate production and wealth accumulation as follows:

$$Y = \int_{-\infty}^{s} \Theta_s x_s^{\gamma} \ell_s^{1-\gamma} ds = \bar{\Theta} X^{\gamma} L^{1-\gamma},$$
$$\dot{X} = \frac{\alpha}{\eta} Y + (r-\delta) X + \underline{x},$$

where  $X = \int_{-\infty}^{t} x_s ds$  and

$$\bar{\Theta}_t = \left(\frac{\int_{-\infty}^s \Theta_s^{1/\gamma} x_s \mathrm{d}s}{\int_{-\infty}^t x_s \mathrm{d}s}\right)^{\gamma}$$

is the wealth-weighted average productivity (which can be taken as exogenous given the evolution of  $\{x_s\}$ ). Therefore, without the cohort-specific labor subsidies, this extension of the model is isomorphic to our baseline model.

### Planner's allocation The planner solves:

$$\max_{\{c,b,\{\ell_s,x_s,\tau_s\}_s\}_t} \int_0^\infty e^{-\rho t} u(c) \mathrm{d}t$$

subject to

$$\mu: \qquad \dot{b} = rb + \int_{-\infty}^{t} \left( (1 - \alpha)y_s(x_s, \ell_s) + \delta x_s - \tau_s \right) \mathrm{d}s - c - \underline{x}_t,$$
  
$$\mu\nu_s: \qquad \dot{x}_s = \frac{\alpha}{\eta}y_s(x_s, \ell_s) + (r - \delta)x_s + \tau_s,$$
  
$$\mu\omega: \qquad L = \int_{-\infty}^{t} \ell_s \mathrm{d}s.$$

Note that we have included transfers  $\tau_s$  for completeness of characterization and will drop them shortly, just after noting that the first-order conditions with respect to  $\tau_s$  simply imply  $\nu_s \equiv 1$  and do not affect any other FOCs.

As usual, optimality conditions with respect to c and b imply

$$u'(c) = \mu = const.$$

We now characterize optimality with respect to  $\{\ell_s, x_s\}$ :

$$\frac{\partial \mathcal{H}}{\partial \ell_s} = \mu \left[ (1-\alpha) + \nu_s \frac{\alpha}{\eta} \right] (1-\gamma) \frac{y_s}{\ell_s} - \mu \omega = 0,$$
  
$$\dot{(\mu\nu_s)} - \rho \mu \nu_s = -\frac{\partial \mathcal{H}}{\partial x_s} = -\mu \left[ (1-\alpha) + \nu_s \frac{\alpha}{\eta} \right] \gamma \frac{y_s}{x_s} - \mu \nu_s (r-\delta) - \mu \delta_s$$

which we rewrite as:

$$(1 + \gamma(\nu_s - 1))(1 - \alpha)\frac{y_s}{\ell_s} = \omega$$

and

$$\dot{\nu}_s = \delta(\nu_s - 1) - \left(1 + \gamma(\nu_s - 1)\right) \frac{\alpha}{\eta} \frac{y_s}{x_s}.$$

Note the difference from our usual  $\dot{\nu}$  equation as we have  $\delta(\nu_s - 1)$  instead of  $\delta\nu_s$  — this is because of the dividend payback. Otherwise everything is identical to our multi-sector tradable model. Denote with  $\varsigma_s = \gamma(\nu_s - 1)$  the sectoral subsidy, so that the planner's allocation is:

$$\ell_s = \frac{[(1+\varsigma_s)\Theta_s]^{1/\gamma} x_s}{\int_{-\infty}^t [(1+\varsigma_{\tilde{s}})\Theta_{\tilde{s}}]^{1/\gamma} x_{\tilde{s}} \mathrm{d}\tilde{s}} L.$$

We have in addition:

$$\dot{\varsigma}_s = \delta \varsigma_s - (1 + \varsigma_s) \gamma \frac{\alpha}{\eta} \Theta_s \left(\frac{\ell_s}{x_s}\right)^{1-\gamma},$$
$$\frac{\dot{x}_s}{x_s} = (r - \delta) + \frac{\alpha}{\eta} \Theta_s \left(\frac{\ell_s}{x_s}\right)^{1-\gamma},$$

which we can rewrite as:

$$\dot{\varsigma}_{s} = \delta\varsigma_{s} - \gamma \frac{\alpha}{\eta} \left[ \frac{(1+\varsigma_{s})\Theta_{s}}{(1+\bar{\varsigma})\bar{\Theta}} \right]^{1/\gamma} \left(\frac{L}{X}\right)^{1-\gamma},$$
$$\dot{\underline{x}}_{s} = (r-\delta) + \frac{\alpha}{\eta} \frac{1}{1+\varsigma_{s}} \left[ \frac{(1+\varsigma_{s})\Theta_{s}}{(1+\bar{\varsigma})\bar{\Theta}} \right]^{1/\gamma} \left(\frac{L}{X}\right)^{1-\gamma}.$$

where

$$(1+\bar{\varsigma_t})\bar{\Theta}_t \equiv \left(\frac{1}{X_t}\int_{-\infty}^t [(1+\varsigma_s)\Theta_s]^{1/\gamma}x_s \mathrm{d}s\right)^{1-\gamma} \quad \text{and} \quad X_t \equiv \int_{-\infty}^t x_s \mathrm{d}s.$$

From this we see that  $\varsigma_s$  is monotonically increasing in  $\Theta_s/\bar{\Theta}$ , and therefore sectors get subsidized based on  $\Theta_s$  and labor allocation is tilted away from being proportional to  $\Theta_s x_s^{\gamma}$  towards being proportional to  $\Theta_s$ . If g = 0, and all  $\Theta_s = const$ , then there are no subsidies and we have a steady state with a lifecycle driven by accumulation of  $x_s$  from  $\underline{x}$  towards a steady state. If g > 0, then new cohorts have  $\Theta_s(t)/\bar{\Theta}(t) > 1$ , and are subsidized initially, and so catch up; the older cohorts are still large, but start to lag behind both due to lower  $\Theta_s/\bar{\Theta} < 1$  and to the relative tax  $(1+\varsigma_s)/(1+\bar{\varsigma}) < 1$ , and eventually vanish, faster under planner's allocation than in the laissez-faire.