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

Oleg Itskhoki and Benjamin Moll
Princeton University

A Derivations and Proofs

Contents

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 ℓ, consider the system of equations

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

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

These two equations define ℓ and c as a function of the marginal utility μ and the wage rate w. The solution for ℓ 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 (μ, w):

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

where the second equality comes from a full differential of (A1)–(A2) under constant μ, 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, ℓ). 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:

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

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 sufficient amount of productivity heterogeneity (η small enough).[1] ■

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

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

[1] 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

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

where we substitute in the policy functions (8)–(9) into the definitions of κ and ℓ, and then took integrals 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:

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

where we isolate out the κ and ℓ 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 aggregate production function to obtain expression (16) for aggregate output y as a function of ℓ and x. The remaining equation 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:

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

A1.3 Entrepreneurs: value and policy functions

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

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

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 $\mathbb{E}\{\mathrm{d}v_t(a,z)\} = B\mathrm{d}a/a + B\mathbb{E}\{\mathrm{d}\tilde{v}_t(z)\}$. Rewrite the value function:

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

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

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

Next we calculate expected value:

$$V_0(a_0) = \int v_0(a_0,z)g_z(z)\mathrm{d}z = \frac{1}{\delta}\big(\tilde{V}_0 + \log a_0\big),$$

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

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

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:

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

and hence

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

We now calculate the average return in our model:

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

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

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

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*

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

holding constant $\ell(t)$ and $c_e(t)$ for all $t \geq 0$. 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 $\{c,\ell,b,x,c_e\}_{t\geq 0}$ satisfying the household and entrepreneurs budget constraints:

$$\dot{b}(t) = (1-\alpha)y\big(x(t),\ell(t)\big) + r^*b(t) - c(t), \tag{A9}$$
$$\dot{x}(t) = \frac{\alpha}{\eta}y\big(x(t),\ell(t)\big) + r^*x(t) - c_e(t), \tag{A10}$$

starting from (b_0, x_0), consider a perturbation $\tilde{x}(t) \equiv x(t) + \beta\hat{x}(t)$, where β 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:

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

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 $t \in (0,t')$:

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

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 $t \le t'$ to obtain:

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

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

$$\hat{x}(t'_-)+\hat{b}(t'_-) = \hat{x}_0 e^{r^* t'}\left[\left(\exp\left\{\gamma\int_0^{t'}\frac{\alpha}{\eta}\frac{y\big(x(t),\ell(t)\big)}{x(t)}\mathrm{d}t\right\}-1\right) + (1-\gamma)\int_0^{t'} e^{-r^* t}\frac{\alpha}{\eta}\frac{y\big(x(t),\ell(t)\big)}{x(t)}\frac{\hat{x}(t)}{\hat{x}_0}\mathrm{d}t\right]$$

where t'_- denotes an instant before t', and we have used the functional form for $y(\cdot)$ and definition of γ 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.[2]

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

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

result 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, ς_x. In particular, the budget constraints of workers, entrepreneurs, and the government (21), (12) and (22) become

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

[2] Note that for small t', we have the following limiting characterization:

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

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 lump-sum transfer which makes the savings decision of entrepreneurs analytically intractable.[3] In what follows we refer to ς_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 τ_ℓ and τ_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 household budget constraint and aggregate wealth accumulation equation.

We impose an additional constraint on the aggregate transfer:[4]

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

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:

$$\begin{aligned} \max_{\substack{\{c,\ell,b,x\}_{t\geq 0},\\ \{\varsigma_x:\ s\leq\varsigma_x x\leq S\}_{t\geq 0}}} \quad & \int_0^\infty e^{-\rho t} u(c,\ell)\mathrm{d}t \\ \text{subject to} \qquad & c + \dot{b} = (1-\alpha)y(x,\ell) + r^* b - \varsigma_x x, \\ & \dot{x} = \frac{\alpha}{\eta} y(x,\ell) + (r^* + \varsigma_x - \delta)x, \end{aligned} \tag{P2}$$

given the initial conditions b_0 and x_0. We still denote the two co-states by μ 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:

$$\nu \geq 1, \quad \varsigma_x x \leq S \qquad \text{and} \qquad \nu \leq 1, \quad \varsigma_x x \geq s. \tag{A12}$$

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,

[3] 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.

[4] 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 (ς_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, $x(t)$.

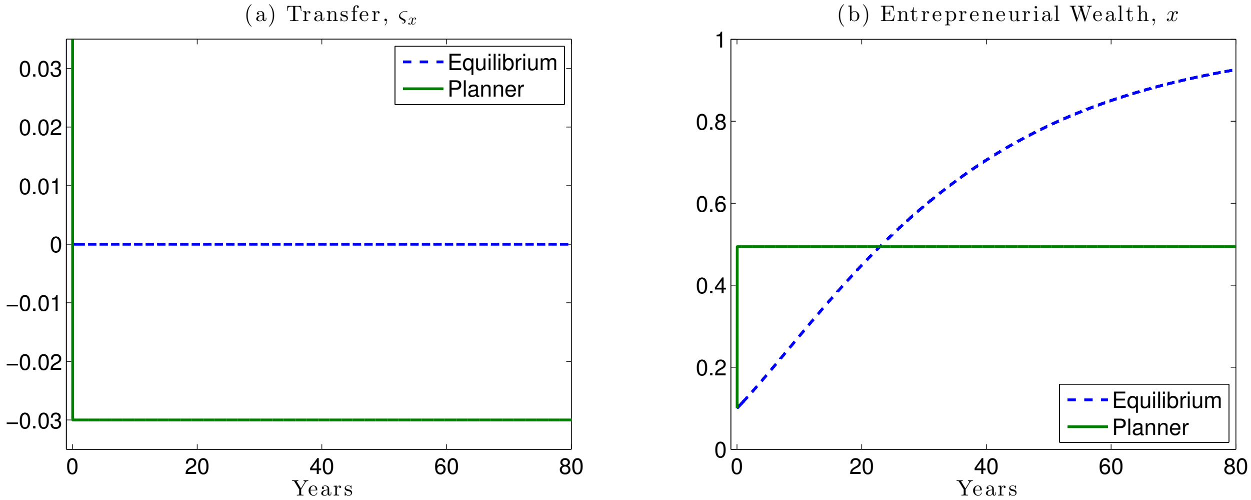

Figure A6: Planner's allocation with unlimited transfers

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

$s < \varsigma_x x < S$, the co-state for the wealth accumulation constraint is unity, $\nu = 1$. In particular, this 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 transfers 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 steady state 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 $x(0+) = \bar{x}$.*[6]

[5]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:

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

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).

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

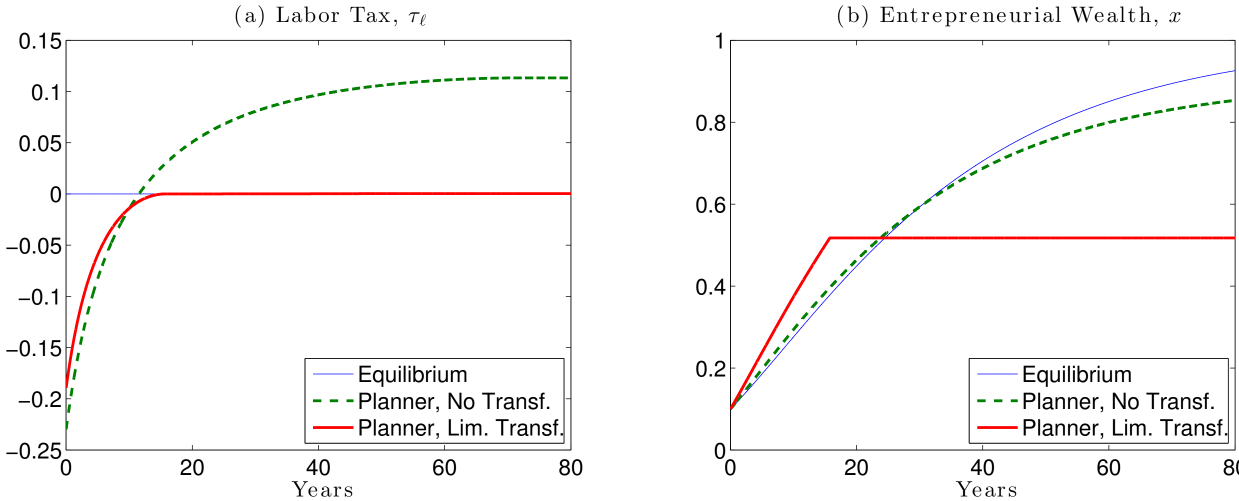

Figure A7: Planner's allocation with limited transfers

Note: evolution of the labor tax $\tau_\ell(t)$ and entrepreneurial wealth $x(t)$ when transfers to entrepreneurs are bounded by $\varsigma_x(t)x(t) \le S < \infty$, while s is not binding (i.e., $s \le -r^*\bar{x}$); 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.

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 between households and entrepreneurs, he 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 Section 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 $s \le -r^*\bar{x}$, 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 $t \ge \bar{t}$. 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 $S < \infty$, $s \le -r^*\bar{x}$, 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 $\big(x(t), \tau_\ell(t)\big)$ 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 $t \ge \bar{t}$, $x(t) = \bar{x}$, $\tau_\ell(t) = 0$ and $\varsigma_x(t) = -r^*$, corresponding to the steady state in Proposition A4. For all $t \ge 0$, $\tau_b(t) = 0$.*

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 steady state 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 agent 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.[7]

[7] For tractability, the way we set up the Ramsey problem without transfers, the subsidy to labor supply is

Furthermore, and as already mention 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 disposable. 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.[8] 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 planner's problem (P1), which allows for (possibly bounded) direct transfers between workers and entrepreneurs (also see Appendix A2.2 below for a more detailed introduction of such transfers). 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 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. We label the resulting planner's problem as (P2), which is explicitly stated in Appendix A2.2, and write the associated present-value Hamiltonian for this problem as:

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

where we have introduced two additional Lagrange multipliers $\mu\bar{\xi}$ and $\mu\underline{\xi}$ for the corresponding bounds on transfers. 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, which we take to be a realistic benchmark.

[8] 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 markets, 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.

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

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

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

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

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

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

$$\bar{\xi} \geq 0, \quad \varsigma_x x \leq S \qquad \text{and} \qquad \underline{\xi} > 0, \quad \varsigma_x x \geq s. \tag{A18}$$

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 $\mu \equiv \bar{\mu}$ and $r^* = \rho$ 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) in the text.

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

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

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

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

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 below 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:

$$\begin{aligned} \max_{\{c,\ell,x,m\}_{t\geq 0}} & \int_0^\infty e^{-\rho t} u(c, 1-\ell)\mathrm{d}t \\ \text{subject to} \quad & \dot{m} = \big(1-\alpha+\alpha/\eta\big) y(x,\ell) + r^* m - \delta x - c, \end{aligned} \tag{P3}$$

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/ℓ 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, τ_ℓ) is 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 $\nu - 1 = -\underline{\xi} < 0$ and $\bar{\xi} = 0$, in which:

$$\frac{\alpha}{\eta}\frac{y\big(\bar{x}, \bar{\ell}\big)}{\bar{x}} = (\delta - r^*) - \frac{s}{\bar{x}} < \delta,$$

$$\bar{\tau}_\ell = \gamma\big(1 - \bar{\nu}\big) = \frac{\gamma}{\gamma + (1-\gamma)\frac{\delta \bar{x}}{r^*\bar{x} + s}} > 0.$$

   When this regime (steady state) is reached, there is a jump from 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: $\underline{\xi} = \bar{\xi} = \bar{\nu} - 1 = \bar{\tau}_\ell = 0$. 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, τ_ℓ 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 ℓ, 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 the bounds follow from (A3). Substituting $\ell(x,\tau_\ell)$ into (29) and (26), we have a system of two autonomous ODEs in (τ_ℓ, x)

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

We now show that the dynamics of this system in (τ_ℓ, 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

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

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

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

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

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

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

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, τ_ℓ) 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 $\dot{x} = 0$ 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, τ_ℓ) 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 19 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 additional subsidies to entrepreneurs: an asset subsidy ς_x, a profit subsidy ς_π, a sales (revenue) subsidy ς_y, a capital subsidy ς_k, and a wagebill subsidy ς_w. Under these circumstances, the budget set of an entrepreneur can be represented as:

$$\dot{a} = (1 + \varsigma_\pi)\pi(a, z) + (r^* + \varsigma_x - \delta)a, \tag{A26}$$

$$\text{with} \qquad \pi(a, z) = \max_{n \geq 0, 0 \leq k \leq \lambda a} \left\{ (1+\varsigma_y) A (zk)^\alpha n^{1-\alpha} - (1-\varsigma_w) wn - (1-\varsigma_k) r^* k \right\},$$

which generalizes expression (33) in the text, and where we already incorporated the optimal consumption-savings decisions of the 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 Θ *and* γ *are defined as in Lemma 2, and we have:*

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

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

$$\begin{aligned}
k &= \lambda a\mathbf{1}_{\{z\geq\underline{z}\}},\\
n &= \left((1-\alpha)\frac{(1+\varsigma^y)A}{(1-\varsigma^w)w}\right)^{1/\alpha} zk,\\
\pi &= \left[\frac{z}{\underline{z}}-1\right](1-\varsigma^k)r^*k,
\end{aligned}$$

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

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

Finally, labor demand in the sector is given by:

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

and aggregate output is given by:

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

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:

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

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 affect 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

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

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 ς_w drops out from the constraints (just like w does in Lemma 2), and it can be recovered from

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

assuming $\tau_\ell=0$, otherwise there is implementational indeterminacy since τ_ℓ and ς_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 Appendix A3.1 in planner's problem (P3) 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:

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

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 ς_π and ς_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 planner's problem (P4) we restrict $\varsigma^x = \varsigma^\pi \equiv 0$, and write the resulting Hamiltonian:

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

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

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

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.[9]

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 ς^k. This last expression characterizes the optimal labor wedge, so that from (A31) we have:

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

complete the proof of the claim in the proposition. ■

[9]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:

$$\tilde{W}_0 = \int_0^\infty e^{-\varrho\tau} q U_0(\tau)\mathrm{d}\tau + \int_{-\infty}^0 q e^{q\tau} \int_0^\infty e^{-(\rho+q)t} u_\tau(t)\mathrm{d}t\mathrm{d}\tau. \tag{A32}$$

This criterion discounts remaining lifetime utility of those currently alive ($\tau \leq 0$) to time $t = 0$ rather than to their birth at time $t = \tau < 0$, or in other words the planner uses the remaining lifetime utility $\tilde{U}_0(\tau) = \int_0^\infty e^{-(\rho+q)t} u_\tau(t)\mathrm{d}t$ for $\tau < 0$. Then the planner integrates remaining lifetime utilities across the living ($\tau < 0$) using their population density $q e^{q\tau}$ and adds the planner-discounted 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 $\rho \leq \varrho \leq \rho + q$ and $\rho = r^*$, and subject to:

$$\begin{aligned} \dot{b} &= r^* b + (1-\alpha) y(x,\ell) - c, \\ \dot{x} &= \frac{\alpha}{\eta} y(x,\ell) + (r^* - \delta)x, \end{aligned}$$

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

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

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:

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

The remaining two conditions characterize the optimal labor wedge:

$$-\frac{u_\ell}{u_c} = \big[1 + \overbrace{\gamma(\nu - 1)}^{=\tau_\ell}\big](1-\alpha)\frac{y}{\ell},$$

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

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 $\varrho > \rho$ chooses a declining path of consumption since $u_c = \mu$ and $\dot{\mu}/\mu = \varrho - \rho > 0$ (i.e., front-loading of consumption to earlier generation 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 ϱ. 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 ν satisfies:

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

where the last term on the right was previously absent. Given that the economy is growing and the planner is impatient, $b \geq 0$ 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 $\varrho \geq \rho$. The more impatient is the planner, the more binding is the constraint, and the larger is ψ. The presence of $\psi > 0$ is equivalent to larger discount rate δ, making the accumulation of wealth x (and its contribution to future productivity) less valuable to the planner. This is a general effect from borrowing constraint 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 ϱ 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 start the transition with a labor subsidy, as in the baseline model. See illustration in Figure A8.

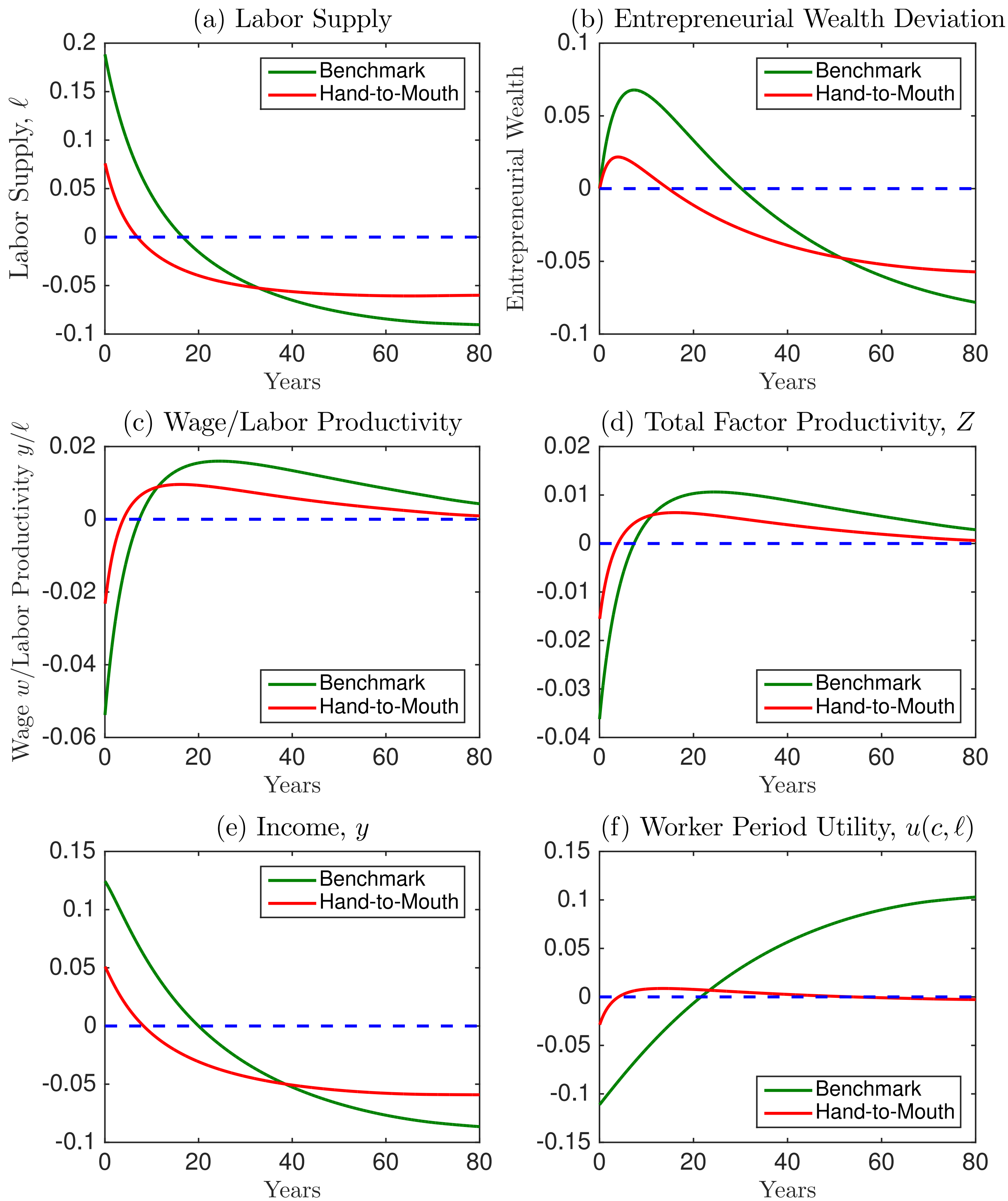

Figure A8: Households with finite lives and borrowing constraints

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 entrepreneurs draw their productivity from a Pareto distribution $G_j(z) = 1 - \big(z/b_j\big)^{-\eta_j}$, where b_j is a lower bound and η_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, at a Poisson rate p (q) the L (H) entrepreneurs becomes H (L) entrepreneurs at any instant (i.e., over any interval of time, the type distribution follows a Markov process).

Note that this way of modeling productivity process maintain the tractability of our framework due to a continuous productivity distribution within types, yet allows us to accommodate arbitrary amount of persistence in the productivity process over time. Indeed, by varying b_j, η_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 $b_H > b_L$.[10] 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 period behavior of entrepreneur is characterized by Lemma 1 independently of the type of the entrepreneur (i.e., independently of whether z was draw 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:

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

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

$$\dot{x}_L = \frac{\alpha}{\eta} y_L(x_L, \ell_L) + (r^* - \delta) x_L + q x_H - p x_L, \tag{A33}$$

$$\dot{x}_H = \frac{\alpha}{\eta} y_H(x_H, \ell_H) + (r^* - \delta) x_H + q x_L - p x_H, \tag{A34}$$

and the labor market clearing requires $\ell_L + \ell_H = \ell$, where ℓ 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

[10] 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:

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

The household budget constraint can then be written as:

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

Following the same steps of 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:

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

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 a weighted average of the two co-states ν_L and ν_H for x_L and x_H respectively, where the weights are the employment shares.[11]

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

$$\begin{aligned}
\dot{\nu}_L &= (\delta+p)\nu_L - q\nu_H - \big[\gamma\nu_L + (1-\gamma)\xi_L\big]\frac{\alpha}{\eta}\frac{y_L(x_L,\ell_L)}{x_L}, \\
\dot{\nu}_H &= (\delta+q)\nu_H - p\nu_L - \big[\gamma\nu_H + (1-\gamma)\xi_H\big]\frac{\alpha}{\eta}\frac{y_H(x_H,\ell_H)}{x_H},
\end{aligned}$$

and (x_L, x_H) are both low, then (ν_L, ν_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 persistent productivity process.

[11] The underlying FOCs are

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

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

A4.2 Pareto weight on entrepreneurs

Consider an extension to the planning problem (P1) in Section 3.2 (without transfers, $\varsigma_x \equiv 0$) in which the planner puts a positive Pareto weight $\theta > 0$ on the utilitarian welfare criterion of all entrepreneurs $\mathbb{V}_0 \equiv \int V_0(a)\mathrm{d}G_{a,0}(a)$ where $V_0(\cdot)$ is the expected value to an entrepreneur with initial assets a_0. From Appendix A1.3, we have:

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

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:

$$\max_{\{c,\ell,b,x\}_{t\geq 0}} \int_0^\infty e^{-\rho t}u(c,\ell)\mathrm{d}t + \frac{\theta}{\delta}\int_0^\infty e^{-\delta t}\frac{\alpha}{\eta}\frac{y\big(x,\ell\big)}{x}\mathrm{d}t \tag{P7}$$

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

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

The Hamiltonian for this problem is:

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

and the optimality conditions are $u_c(t) = \mu(t) = \bar{\mu}$ 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,ν) is the same as in Section 3.2 with the exception of an additional term $\frac{\theta}{\delta\bar{\mu}}e^{-(\delta-\rho)t}\frac{\gamma}{x} \geq 0$ in the condition above. Similarly, the optimal labor wedge which we denote by τ_ℓ^θ is given by (32). Note that the long-run optimal tax rate is the same for all $\theta \geq 0$, as a consequence of our assumption that entrepreneurs are more impatient than workers, $\delta > \rho$. When $\delta = \rho$, the long-run tax depends on θ and can be negative for θ 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.[12] In what follows, we set up formally the closed economy model. In particular, we generalize Lemmas 2 and 3

[12]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, 2013).

to show that the constraints on allocations (25)–(26) in the closed economy become:

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

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

and where the output function is now:

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

instead of (16). 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, κ.

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 worker's savings in the early phase of transition, provided x/κ is low enough. This allows the economy to accumulate capital, κ, 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/κ is in the steady state. The qualitative prediction 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.[13]

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 in capital demand (13) into the aggregate production function (15), we obtain (A40) which defines $y(x, \kappa, \ell)$ in the text. We can then summarized the planner's problem in the closed economy as:

$$\max_{\{c,\ell,\kappa,b,x\}} \int_0^{\infty} e^{-\rho t} u(c, \ell) \mathrm{d}t, \tag{PC}$$

$$\text{subject to} \quad \dot{b} = \left[(1-\alpha) + \alpha\frac{\eta-1}{\eta}\frac{b}{\kappa}\right] y(x, \kappa, \ell) - c,$$

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

and given (b_0, x_0), $\kappa = x + b$, and where we have used (18) to substitute out endogenous interest rate r.

To simplify notation, we replace the first constraint with the sum of the two constants to

[13]Formally, we show that, in the absence of transfers, the optimal tax on labor supply and worker savings satisfy:

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

where $\nu(t)$ is again the co-state for $x(t)$.

substitute κ for $b + x$. The Hamiltonian for this problem is:

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

and the optimality conditions are:

$$\begin{aligned}
0 = \frac{\partial\mathcal{H}}{\partial c} &= u_c - \mu,\\
0 = \frac{\partial\mathcal{H}}{\partial \ell} &= u_\ell + \mu\left[1 + \nu\frac{\alpha}{\eta}\left(1 + (\eta-1)\frac{x}{\kappa}\right)\right](1-\alpha)\frac{y}{\ell},\\
\dot{\mu} - \rho\mu = -\frac{\partial\mathcal{H}}{\partial \kappa} &= -\mu r - \mu\nu\frac{\alpha}{\eta}\left(1 + (\eta-1)\frac{x}{\kappa}\right)r + \mu\nu\frac{x}{\kappa}r,\\
\dot{(\mu\nu)} - \rho\mu\nu = -\frac{\partial\mathcal{H}}{\partial x} &= -\mu\left(\frac{\alpha}{\eta}\frac{y}{x} - \delta\right) - \mu\nu\left(\frac{\alpha}{\eta}\left(1 + (\eta-1)\frac{x}{\kappa}\right)\frac{\alpha}{\eta}\frac{y}{x} + r - \delta\right).
\end{aligned}$$

From the second condition we have labor wedge:

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

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

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

Finally, we have:

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

This dynamic system can be solved using conventional methods. Note that ν in this problem corresponds to $(\nu - 1)$ in the text, as we have used the sum of the two constraint (country aggregate resource constraint) instead of using the household budget constraint.

A4.4 Optimal intertemporal wedge

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

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

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:

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

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

$$\begin{aligned}
0 &= \frac{\partial\mathcal{H}}{\partial c} = u_c - \mu\big(1+\psi h_c\big),\\
0 &= \frac{\partial\mathcal{H}}{\partial \ell} = u_\ell + \mu\big(1-\gamma+\gamma\nu\big)(1-\alpha)\frac{y}{\ell} + \mu\psi\left((1-\gamma)(1-\alpha)\frac{y}{\ell} - h_\ell\right),\\
\dot{\mu} - \rho\mu &= -\frac{\partial\mathcal{H}}{\partial b} = -\mu r^*,\\
\dot{(\mu\nu)} - \rho\mu\nu &= -\frac{\partial\mathcal{H}}{\partial x} = -\mu\big(1-\gamma+\gamma\nu\big)\frac{\alpha}{\eta}\frac{y}{x} - \mu\nu(r^*-\delta) - \mu\psi(1-\gamma)\frac{\alpha}{\eta}\frac{y}{x}.
\end{aligned}$$

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:

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

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

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

Finally, the dynamics of ν satisfies:

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

and the distortion to the consumption smoothing satisfies:

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

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.

A5 Appendix for Quantitative Model in Section 4

This Appendix describes 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 instrument. The generalization to other tax instruments is straightforward, and we detail the specific case of the 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:

$$c(t) = (1 - \tau_\ell(t))w(t)\ell(t) + T(t). \tag{A43}$$

Entrepreneurs As explained in the main text, it now becomes necessary to keep track of the joint distribution of entrepreneur 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:

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

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).[14] With this notation in hand, the system of PDEs summarizing entrepreneurs' behavior is:

$$\begin{aligned} \rho v(a, z, t) = &\max_c \; u(c) + \partial_a v(a, z, t)[\pi(a, z; w(t), r^*) + r^* a - c] \\ &+ \partial_z v(a, z, t)\mu(z) + \frac{1}{2}\partial_{zz} v(a, z, t)\sigma^2(z) \\ &+ \phi \int_{\underline{z}}^{\bar{z}} (v(a, x, t) - v(a, z, t))\, p(x) dx + \partial_t v(a, z, t) \end{aligned} \tag{A45}$$

$$\begin{aligned} \partial_t g(a, z, t) = &- \partial_a[s(a, z, t)g(a, z, t)] - \partial_z[\mu(z)g(a, z, t)] + \frac{1}{2}\partial_{zz}[\sigma^2(z)g(a, z, t)] \\ &- \phi g(a, z, t) + \phi p(z) \int_{\underline{z}}^{\bar{z}} g(a, x, t) dx, \end{aligned} \tag{A46}$$

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

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_∞ 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 $\bar{z}$, namely $\partial_z v(a, \underline{z}, t) = \partial_z v(a, \bar{z}, t) = 0$ for all (a, t). For numerical reasons, we also impose a state constraint $a \geq 0$ 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 ϕ, 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 $\int_{\underline{z}}^{\bar{z}} g(a, x, t)dx$) switch to productivity type z

[14] 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 $\mathcal{A}f(z) = \mu(z)f'(z) + \frac{1}{2}\sigma^2(z)f''(z) + \phi \int_{\underline{z}}^{\bar{z}} (f(x) - f(z))p(x)dx$, 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 η. The boundary conditions are those corresponding to reflecting barriers: $f'(\underline{z}) = f'(\bar{z}) = 0$. 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 $(-\nu \log z_t + \sigma^2/2)z_t$.

at rate $\phi p(z)$ and hence the inflow term $+\phi p(z) \int_{\underline{z}}^{\bar{z}} g(a,x,t)dx$.[15]

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:

$$\omega \int_{\underline{z}}^{\bar{z}} \int_0^\infty n(a,z;w(t),r^*) g(a,z,t) da dz = (1-\omega)\ell(t). \tag{A49}$$

Given initial condition $g_0(a,z)$, the two PDEs (A45), (A46) together with (A47), workers' optimality condition (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 do the grid search.

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 Δ. Denoting the equilibrium allocation under laissez-faire with hats, Δ solves

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

where recall that the $t=0$ value to entrepreneur is $v_0(a,z) = \mathbb{E}_0 \left[\int_0^\infty e^{-\delta t} \log(c^e_t) \mathrm{d}t \middle| (a_0,z_0) = (a,z) \right]$.

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

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

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

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

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 $\mathcal{V}_0^w = \int_0^\infty e^{-\rho t} u(c_t, \ell_t)\mathrm{d}t$ and $\mathcal{V}_0^e = \int_{\underline{z}}^{\bar{z}} \int_0^\infty v_0(a,z) g_0(a,z) da dz$ 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.

Table A2: Parameter Values for Quantitative Exercise

Param.	Value	Description	Comment/Source
χ	0.1	scaling factor of initial dist.	initial =1/10 stationary wealth
λ	2	tightness of financial constraint	Steady state $D/Y = 2.29$
α	0.33	capital share	standard value
β	0.90	returns to scale	De Loecker et al. (2016)
$e^{-\nu}$	0.85	autocorrelation of productivity	Asker et al. (2014)
σ	0.3	innovation variance of productivity	Asker et al. (2014)
ϕ	0.1	arrival rate of Poisson shocks	jump on average every 10 yrs
ζ	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^*
δ	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

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 χ of aggregate final wealth $\int a dG_0(a,z) = \int a dG_\infty(a/\chi, z) = \chi \int a d\mathcal{G}_\infty(a,z)$. We then set $\chi = 0.1$. The capital share α and returns to scale β are set to standard parameter values from 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 low value of the returns-to-scale parameter β. 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 β 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).[16] Also as discussed in the text, λ 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 $Y_t = \int y_t(a,z) d\mathcal{G}_t(a,z)$.

As in the baseline model, we set workers' discount rate ρ 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 δ 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 ℓ 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:

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

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:

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

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

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

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

[16] See their Online Appendix Table OA.11 at http://www.princeton.edu/~jdeloeck/ACWDLappendix.pdf

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 wage rate and the labor share of the unconstrained entrepreneurs, $\{\hat{\tau}_\ell^U(t), w(t), \pi(t)\}_{t\geq 0}$. For a given path of $\{\hat{\tau}_\ell^U(t), w(t)\}$ 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:

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

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

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

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

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

Analogously to above, we search for a triple of parameters $(\underline{\varsigma}_k, \bar{\varsigma}_k, \gamma_k)$ 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. Later, we discuss 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 $i = 1, ..., k$ are tradable and goods $i = k+1, ..., n$ are not tradable. The households maximize the intertemporal utility given by $\int_0^\infty e^{-\rho t} u(t)\mathrm{d}t$, and supply inelastically a total of L units of labor,

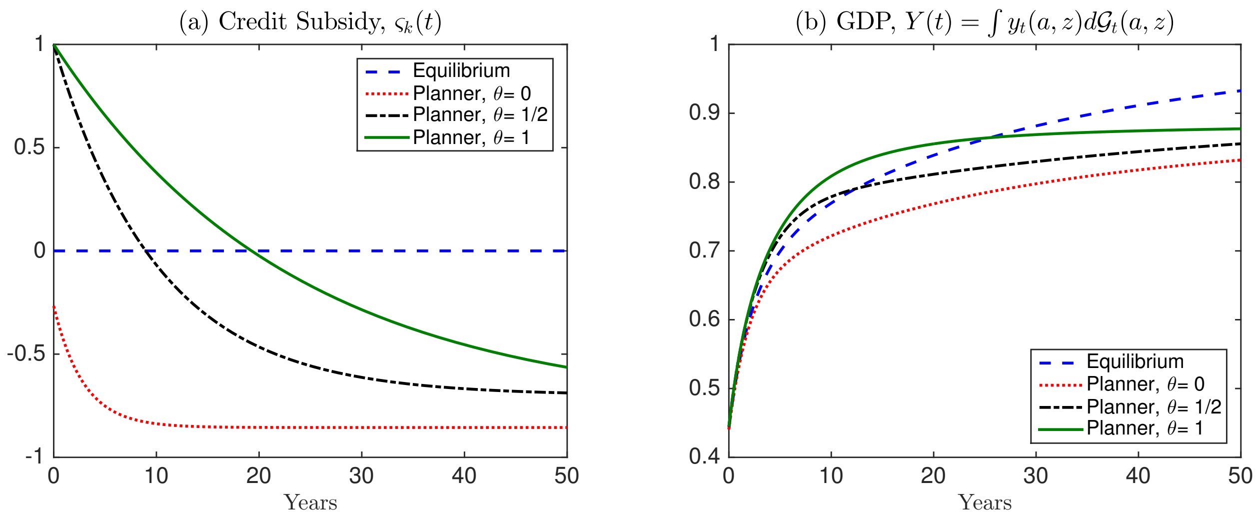

Figure A9: Optimal policy in quantitative model: credit subsidy $\varsigma_k(t)$ and GDP $Y(t)$

Note: steady state GDP in the laissez-faire equilibrium is normalized to 1 in panel (b).

which is split between the sectors:[17]

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

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

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

where w is the common after-tax wage, w_i is the wage paid by the firms in sector i and τ_i^ℓ is the tax on labor income earned in sector i.[18]

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:

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

where p_i is producer price of and τ_i^c is the consumption tax on good i, b is the asset position of the households and τ^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:

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

where $u_i \equiv \partial u / \partial c_i$ is the marginal utility from consumption of good i. The first condition is the

[17]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.

[18]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.

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.

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 λ_i, operate sector-specific Cobb-Douglas technologies with productivity A_i and capital-intensity α_i, and draw their idiosyncratic productivities from sector-specific Pareto distributions with tail parameter η_i. As a result, and analogous to Lemma 2, the aggregate production function in sector i is:[19]

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

where

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

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 κ_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 δx_i, where δ is the entrepreneurial discount rate.[20] 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 α_i/η_i is the share of profits in the sectoral revenues, and $(1 - \alpha_i)$ is the share of labor income:

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

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

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

Note that we rule out direct sectoral transfers which would allow the planner to effectively sidestep the financial constraints.[21]

[19] 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 ς_y in that derivation.

[20] We make this assumption for tractability, but the analysis extends to more general utility functions of entrepreneurs.

[21] 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 instruments is used, just as in the one-sector economy in Appendix A2.2.

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 $(p_0, p_1, \dots, p_k)$ for $k \le n$. The prices of the non-tradables p_i for $i = k+1, \dots, n$ are determined to clear the respective markets:

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

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 $\big(\{c_i, \ell_i, x_i\}_{i=0}^N, b\big)_{t\ge 0}$ that satisfies the following dynamic system:*

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

there exists a path of taxes $(\tau^b, \{\tau_i^c, \tau_i^\ell\}, T)_{t\ge 0}$ that decentralize this allocation as an equilibrium in the multi-sector economy, where r^* and $(p_0, \dots, p_k)$ are international prices and $(p_{k+1}, \dots, p_N)$ 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

$$\big(b, \{c_i, \ell_i, x_i\}_{i=0}^N, \{p_i\}_{i=k+1}^N\big)_{t\ge 0}$$

and subject to the set of constraint in (A59) with corresponding Lagrange multipliers denoted by $\mu \cdot \big(1, \{\nu_i\}_{i=0}^N, \{\psi_i\}_{i=k+1}^N, \omega\big)$. Given a dynamic allocation, we recover the corresponding paths of taxes τ^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 ($\tau_i^\ell \equiv 0$) by imposing

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

in addition to constraints in (A59), and we denote the corresponding Lagrange multipliers by $\mu \xi_i$ for $i = 0, \dots, N$.[22] Second, we also rule out static consumption taxes ($\tau_i^c \equiv 0$), leaving the planner

[22] Note that in this case the common wage rate w becomes a variable of planner's optimization.

with only the intertemporal tax τ^b. In this case, we additionally impose

$$u_i = p_i u_0, \qquad i = 1, \ldots, N, \tag{A61}$$

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 for derivation):

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.[23] 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 wide-spread 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:*

$$\tau_i^c = \begin{cases} 0, & i = 0, 1, \ldots, k, \\ \frac{1}{\eta_i - 1}(1 - \nu_i), & i = k+1, \ldots, N, \end{cases} \qquad \text{and} \qquad \tau_i^\ell = \begin{cases} \gamma_i(1 - \nu_i), & i = 0, 1, \ldots, k, \\ -\tau_i^c, & i = k+1, \ldots, N, \end{cases}$$

where ν_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 η_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 labor wedge introduced by the consumption subsidy, $\tau_i^\ell = -\tau_i^c$.

To understand the overall effect of these various tax instruments, it is useful to define the *overall*

[23] The situation is different if the planner has access to production or export taxes for tradable goods, which we discuss below.

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

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

and for non-tradable sectors

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

Consider first the overall labor wedge for non-tradables in (A63). Somewhat counterintuitively, 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, τ_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 price 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 ν_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 ν_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, Θ_i may differ across sectors due to either physical productivity A_i or financial constraints λ_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 in Appendix A6 that a 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 laissez-faire in Figure 5. 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 α_i and η_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.[24] 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

[24]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.

pattern of specialization.[25]

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 $i = 0$ and 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 CPI-based 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.[26] 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 $\alpha_1 \to 0$, 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

[25] In the long-run, somewhat counterintuitively, the planner drives the economy toward specialization in the sector with the *lower* γ_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 γ. This is because a higher γ 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 γ_i, thereby specializing against it in the long-run. Things may be different during the transition.

[26] 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 normalized $p_0 = 1$ 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.

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 allocated jobs and occupation 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 multisector 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, τ_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 planner only uses the non-tradable consumption tax and sets it according to:

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

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 κ captures the fact that the planner uses the consumption tax to also affect the labor allocation. Note that κ increases in ν_0 and decreases in ν_1. Intuitively, if ν_0 is large, then the only way to improve the allocation is by taxing non-tradable consumption (which is reflected in the κ term in τ_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 ν_0, the smaller is ν_1, the larger is the size of the tradable sector 0 (in terms of labor allocated to this sector), and the smaller is γ_1. In particular, as $\gamma_1 \to 0$ (for example, due to $\alpha_1 \to 0$, 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 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, 2012, for the equivalence result of these policies). We provide a formal analysis of this case in Appendix A6, 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 ν_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, price of non-tradables, and consumption of both goods decrease, while the tradable sector expands production and exports, facing unchanged international prices.[27] 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, lthereby 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 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. This contrasts with the narrative in the optimal exchange rate policy literature (see Rodrik, 2008; Korinek and Serven, 2010; Benigno and Fornaro, 2012), which tends to focus on the 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:

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

with $p_0 \equiv 1$, $i \le k$ tradable goods (p_i given exogenously) and $i > k$ non-tradable goods (p_i determined in equilibrium to clear the good market). The constraints with co-state variables μ 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 $\mu\xi_i$

[27] 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.

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 τ_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 τ_i^ℓ are available and unconstrained, we have $\xi_i \equiv 0$ (follows from the FOC for τ_i^ℓ), and hence we simply drop the last line of constraints.
3. With transfers, i.e. with all instruments. Just like in 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 in the very end.

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's are

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

where we have introduced $\psi_i \equiv 0$ for $i \le k$. 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 $i \le k$.

The static FOCs for ℓ_i for all i and p_i for $i > k$ are:

$$\left((1-\alpha_i)(1+\xi_i) + \frac{\alpha_i}{\eta_i}\nu_i + \tau_i^c\right)(1-\gamma_i)\frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega + \xi_i \frac{w}{1-\tau_i^\ell}, \qquad \forall i$$

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

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

$$\left(1 + \frac{1-\gamma_i}{1-\alpha_i}\tau_i^c\right)(1-\alpha_i)\frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega, \qquad \forall i$$

$$\underbrace{\left(1 - \alpha_i + \frac{\alpha_i}{\eta_i}\right)\big[1 + \gamma_i(\eta_i - 1)\big]}_{=1 \text{ from definition of } \gamma_i} + \gamma_i(\eta_i - 1)\tau_i^c = 1, \qquad i > k.$$

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 not $\nu_i = 1$. Therefore, the two sets of conditions are:

$$\left(1 + \gamma_i(\nu_i - 1) + \frac{1-\gamma_i}{1-\alpha_i}\tau_i^c\right)(1-\alpha_i)\frac{p_i y_i(x_i, \ell_i; p_i)}{\ell_i} = \omega, \qquad \forall i$$

$$1 + \gamma_i(\nu_i - 1) + \gamma_i(\eta_i - 1)\tau_i^c = 1, \qquad i > k,$$

where we use the property that:

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

We simplify:

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

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:

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

This complete the proofs of Lemma A7 and Proposition A6 in the text.

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:

$$\begin{aligned}
(1-\gamma_i)(1+\xi_i) + \gamma_i\nu_i + \frac{1-\alpha_i}{1-\gamma_i}\tau_i^c &= 1 + \kappa + \xi_i, \qquad \forall i \\
(1-\gamma_i)(1+\xi_i) + \gamma_i\nu_i + \gamma_i(\eta_i - 1)\tau_i^c &= 1, \qquad i > k, \\
\sum_{i=1}^N \xi_i \ell_i &= 0,
\end{aligned}$$

where in the first set of FOCs we used that $(1-\alpha_i)p_i y_i/\ell_i = w$ 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 ξ_i:

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

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, consumption tax on non-tradables decreases in ν_i and increases in κ, which measures the average scarcity of capital across sectors. Derive the expression for κ 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, κ increases in ν_i for $i \le k$ and decreases in ν_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 refers 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 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 ω, which is a function of time, and where τ_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 ω:

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

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 applied, but with $\tau_i \equiv 0$. The dynamics for this system is 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 τ_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 homogenous good but multiple cohorts producing it. Any living firm dies at a Poisson rate δ and pays out its wealth back to households, while a new inflow of firms happens at rate δ endowed with exogenous wealth $\underline{x}_t$. The productivity of new cohorts increases at rage g: $\Theta_t = \Theta_0 e^{gt}$. We write the production function of cohort $s \in (-\infty, t]$ as:

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

The 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, the 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 ω is the common wage rate. Therefore, ℓ_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 δx_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:

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

Decentralized Allocation There are only two dimensions of choice: 1) intertemporal allocation of consumption and 2) allocation of labor across cohorts $\{\ell_s\}$. The first one is frictionless and irrelevant, 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 the sufficient statistic for labor allocation.

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

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

where $X = \int_{-\infty}^{t} x_s \mathrm{d}s$ 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 takes 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

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

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

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

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

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

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

which we rewrite as:

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

and

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

Note the difference from our usual $\dot{\nu}$ equation that we have $\delta(\nu_s - 1)$ instead of $\delta\nu_s$ — this is because of the dividend pay back. 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:

$$\dot{\varsigma}_s = \delta\varsigma_s - \gamma\frac{\alpha}{\eta}\left[\frac{(1+\varsigma_s)\Theta_s}{(1+\bar{\varsigma})\bar{\Theta}}\right]^{1/\gamma}\left(\frac{L}{X}\right)^{1-\gamma},$$
$$\frac{\dot{x}_s}{x_s} = (r-\delta) + \frac{\alpha}{\eta}\frac{1}{1+\varsigma_s}\left[\frac{(1+\varsigma_s)\Theta_s}{(1+\bar{\varsigma})\bar{\Theta}}\right]^{1/\gamma}\left(\frac{L}{X}\right)^{1-\gamma}$$

where

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

From this we see that ς_s is monotonically increasing in $\Theta_s/\bar{\Theta}$, and therefore sectors get subsidized based on Θ_s and labor allocation is tilted away from being proportional to $\Theta_s x_s^\gamma$ towards being proportional to Θ_s. If $g = 0$, and all $\Theta_s = const$, then there are no subsidies and we have a steady state with a life cycle 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 in finite time.