Estate Taxation and Human Capital with Information Externalities

Aaron Hedlund*
University of Missouri

April 1, 2016

Abstract

This paper investigates the effects of estate taxation when firms cannot directly observe worker skill levels. Imperfect labor market signaling gives rise to an information externality that causes workers to free-ride off of others' human capital acquisition. Inherited wealth exacerbates the information externality because risk averse workers with larger inheritances exert less effort to acquire skills. By reducing these inheritances, an estate tax induces greater skill acquisition effort and increases the number of skilled workers. In a quantitative model with employer learning and capital accumulation, the optimal estate tax is significantly above zero, increases wages and output, and benefits a large majority of households.

Keywords: Information externalities; signaling; free-rider problem; labor markets; bequests; estate taxes

JEL Classification Numbers: D62, D82, E21, E24, E60, H21

1 Introduction

Over a century ago, Andrew Carnegie observed that "the parent who leaves his son enormous wealth generally deadens the talents and energies of the son and tempts him to lead a less useful and less worthy life than he otherwise would." More recently, Warren Buffett has echoed this sentiment by stating his intention not to leave a large fortune to his children. While the habits of heirs and heiresses may appear to be little more than a family or tabloid affair, policymakers are increasingly concerned about growing skill mismatch in the labor

*Comments are welcome at *hedlunda@missouri.edu*. I thank Dirk Krueger, Guido Menzio, Harold Cole, and Andy Postlewaite for many useful comments. Any errors are my own.

market and its implications for unemployment and growth. Therefore, because of how they impact human capital decisions, bequests have potential adverse macroeconomic spillovers.

In this paper, I investigate the consequences of bequests for skill acquisition and macroeconomic performance, and I assess how estate taxation can potentially mitigate negative externalities from inefficient wealth accumulation. To do so, I construct a theory of stochastic skill acquisition, imperfect labor market signaling, and inheritance heterogeneity. In a one period setting, I demonstrate theoretically that when firms cannot perfectly observe worker skills, an information externality arises that causes workers to free-ride off of others' skill acquisition, which depresses the fraction of skilled workers in the economy and reduces wages and output. With risk-averse workers, inheritances magnify the information externality because free-riding increases with inherited wealth. An estate tax can mitigate this externality and induce higher effort, thereby yielding a more skilled workforce, higher wages, and higher output. The estate tax is progressive: low-wealth households experience the greatest welfare gains, while those with the largest inheritances gain the least (or lose).

To understand the economic mechanism behind these theoretical results, it is instructive to explore the nature of the information externality. In the model, skilled and unskilled workers differ in their productivity to firms, but firms only observe noisy signals of each worker's skill level. However, firms are also informed about the overall proportion of skilled workers in the population. Firms form beliefs about the skill level of each worker based on the worker's signal and this proportion. As a result, firms offer signal-specific wages that increase in the population share of skilled workers.

Workers' incentive to acquire skills comes from the fact that skilled workers are more likely to send positive signals to the labor market and, thus, to receive higher wages. The information externality arises because workers benefit from a higher overall proportion of skilled workers but do not internalize how their own skill acquisition affects this proportion. Instead, workers free-ride off of the skill acquisition of others. In the one period model, estate taxation reduces inherited wealth, increases workers' marginal utility of consumption and, thus, the gain to acquiring skills, and partially mitigates the information externality.

Next, I evaluate the quantitative implications of this information externality—along with the resulting inefficient wealth accumulation—and the effect of estate taxation using a dynamic, incomplete markets model that features employer learning and capital accumulation. In the model, youth receive their inheritance and make a one-time decision regarding how much effort to exert attempting to acquire skills. Those who succeed become skilled workers, whereas those who do not become unskilled workers. Once in the labor market, workers send imprecise signals each period to employers, who continually update their beliefs about the skill level of each worker and adjust compensation accordingly. However, in the event of an involuntary job separation, the worker's history of signals at their current employer is lost, and the learning process starts over at a new firm. As in the static model, firms supplement the private signals of new workers with the public signal—the population fraction of skilled workers—to form initial skill beliefs. However, even with learning, the impact of the public signal has persistent effects on history-dependent wages.

The estate tax has a similar effect as in the static model by inducing higher effort, raising the fraction of skilled workers, and increasing the public signal. Furthermore, because stochastic signals and job separations create uninsurable risk in the presence of incomplete markets, estate taxation provides social insurance for workers. However, estate taxation also blunts the incentive to accrue wealth and impacts capital accumulation. A higher skilled worker population improves the public signal and raises wages, but a lower capital stock reduces wages. The theoretical impact of estate taxes on wages in the dynamic model is therefore ambiguous and depends on the magnitudes of these two channels.

To provide a quantitative assessment, the model is calibrated to match the U.S. capital to output ratio, the share of college graduates, and the college wage premium. A central lesson that emerges from the baseline model is the importance of how the government allocates the revenues from the estate tax. When the government uses the revenues for unproductive government spending, implementation of a 50% estate tax raises the share of skilled workers from 30% to almost 36%, but it also decreases the capital stock and has a net negative effect on output, wages, and welfare (in a utilitarian sense—some households still gain). Rebating the revenues lump sum compensates households for the lost wealth, but this mean-preserving compression of the asset distribution actually *reduces* aggregate effort and, with it, the skilled worker population because of the decreasing and convex relationship between individual effort and assets. However, when the government channels the revenue to fund targeted subsidies of new workers who receive high signals—in effect, subsidizing skill acquisition—the 50% estate tax increases the skilled worker population from 30% to over 40%. Furthermore, output increases by 2%, starting wages improve, and welfare jumps by 2.39%. I also show that the benefits of estate taxation are robust but decreasing in signal accuracy and that removing the distortion on capital accumulation greatly increases the positive impact of estate taxation on wages and welfare. Lastly, augmenting the model to include a Lucas (1988) human capital externality enhances the welfare gains of estate taxation but blunts its ability to increase the share of skilled workers in the economy.

Lastly, I determine the optimal estate tax rate in the quantitative model under a variety of specifications. Specifically, I assume that the social planner maximizes the average consumption-equivalent welfare gain for workers and youth of switching from a 0% estate tax to each possible choice of rate $\tau_a \in (0, 1]$. Importantly, the planner takes into account the dynamic transition path of the economy from the initial steady state to the new equilibrium. With this utilitarian objective, the optimal estate tax ranges from 5% with highly accurate labor market signaling to 95% with very noisy signaling. In the baseline specification, the optimal rate is 30%, with 79% of households experiencing welfare gains. Adding Lucas human capital externalities to the baseline model pushes the optimal rate up to 70% and creates substantially larger gains in welfare and output.

1.1 Related Literature

This paper is related to several strands of economic literature. First, the information externality in this paper mirrors the one studied by Fang and Norman (2006), who investigate the effects of government-mandated discrimination in public sector hiring. They show that, under certain conditions, preferential treatment can actually *hurt* those it intends to benefit by dampening the incentive to acquire skills and worsening the information externality. Using a similar information externality, Popov and Bernhardt (2012) show that fraternity membership affects labor market outcomes when firms cannot perfectly evaluate the productivities of job applicants. In the same vein, Lockwood (1991) shows that an information externality arises when firms administer tests that imperfectly measure worker productivity to make hiring decisions. My paper builds upon this literature by examining the relationship between information externalities and wealth and by introducing dynamics with learning.

More broadly, this paper fits into an extensive literature on human capital externalities. As emphasized by both Acemoglu (1996) and Lucas (1988), the rate of return of human capital for individual workers increases in the aggregate human stock. Lucas (1988) focuses on the case of technological increasing returns, whereas Acemoglu (1996) highlights a pecuniary externality that arises from the interaction of ex ante investments and random bilateral job search. Restrepo (2015) takes a similar approach to analyze skill mismatch, job polarization, and structural unemployment. Because of matching frictions, firms cannot perfectly locate workers with the requisite skills to fill novel jobs. This inability to direct search toward recruiting skilled workers causes job creation to depend on the skill composition of the unemployed worker pool, and the availability of jobs impacts workers' incentives to retrain. On the empirical side, Rauch (1993), Iranzo and Peri (2009), Gennaioli, La Porta, Lopez-de Silanes and Shleifer (2013), Lange and Topel (2006), Acemoglu and Angrist (2000), Bedard (2001), and others find widespread and compelling evidence that worker education levels impact total factor productivity and wages, even after controlling for individual characteristics.

Lastly, this paper interfaces closely with the literature on inheritances and estate taxation. Empirically, Holtz-Eakin, Joulfaian and Rosen (1993), Joulfaian and Wilhelm (1994), Brown, Coile and Weisbenner (2010), and Elinder, Erixson and Ohlsson (2012) find that large inheritances have a sizable, negative effect on labor market behavior. More recently, Bø, Halvorsen and Thoresen (2015) use administrative data to uncover significant cross-sectional heterogeneity of the Carnegie effect. In the taxation literature, recent work by Golosov, Kocherlakota and Tsyvinski (2003), Farhi and Werning (2012), Straub and Werning (2015), and others has challenged the conventional Chamley (1986) and Judd (1985) findings of zero long-run optimal capital taxes. In particular, Farhi and Werning (2010) and Piketty and Saez (2013) show the optimality of inheritance taxes in models with altruistic bequests and inequality. Furthermore, Pestieau and Sato (2008), Kapicka and Neira (2013), and Krueger and Ludwig (2016) assess how skill formation and education alter the effects of taxation. My paper contributes to the taxation literature by adding human capital spillovers through an information externality and showing significant *quantitative* positive effects of estate taxes on welfare, wages, and output.

2 The Static Model

2.1 Households

A continuum of households $i \in [0,1]$ draw inheritances $a(i)$ from a cumulative distribution function $\Omega(a)$ with compact support $[0, \overline{a}]$ and continuous density function $f_\Omega(a)$. Households also differ in their degree of labor market skills, $x(i) \in \{U, S\}$. All workers inelastically supply one unit of time to the labor market, but only skilled workers are productive. Specifically, each skilled worker supplies $n_S(i) = 1$ units of effective labor, while each unskilled worker supplies $n_U(i) = 0$ units of effective labor.

At the beginning of the period, workers decide how much effort $e \in [0,1]$ to exert attempting to acquire skills. Effort affects the probability of skill acquisition according to a Bernoulli distribution,

$$x(i) = \begin{cases} S & \text{with probability } e(i) \\ U & \text{with probability } 1 - e(i). \end{cases}$$

Household preferences over consumption c and effort e are given by

$$U(c,e) = u(c) - v(e).$$

Workers' skill levels are private information. Upon entering the labor market, each worker sends an imprecise signal $s(i) \in \{l, h\}$ of its skill level to the labor market, with signal accuracy p satisfying

$$p = \mathbb{P}(s(i) = h | x(i) = S) = \mathbb{P}(s(i) = l | x(i) = U) > \frac{1}{2}.$$

2.2 Firms

Firms use effective labor N to produce output Y according to

$$Y = AN.$$

Firms cannot identify the skill level of workers. Instead, they engage in Bertrand competition to hire workers based on signal. Wages $w_s(\pi)$ equal the expected marginal product of each worker conditional on signal $s \in \{l, h\}$ and the commonly known fraction π of skilled workers in the population.

2.3 Decision Problems

At the beginning of the period, workers make their effort choice, knowing their inheritance a, the proportional estate tax rate τ, and the fraction π of skilled workers in the population. After effort choices are made, workers discover whether they have acquired skills. Workers then enter the labor market, send their signals to employers, and receive their wages after production occurs. The government uses the estate tax revenue for wasteful spending.

2.3.1 Household's Problem

A household with assets a chooses effort e to solve

$$\begin{aligned} V(a; \pi, \tau) = \max_{e \in [0,1]} & [ep + (1-e)(1-p)]u(w_h(\pi) + a(1-\tau)) \\ & + [e(1-p) + (1-e)p]u(w_l(\pi) + a(1-\tau)) - v(e) \end{aligned} \tag{1}$$

Let $e(a; \pi, \tau)$ denote the household's effort policy function. Lemma 1 in the next section establishes the continuity of $e(a; \pi, \tau)$ in both a and π.

2.3.2 Firm's Problem

Firms' beliefs about the skill level of a worker with signal $s(i)$ are given by

$$\mathbb{P}(x(i) = S | s(i) = h) = \frac{\pi p}{\pi p + (1-\pi)(1-p)},$$
$$\mathbb{P}(x(i) = S | s(i) = l) = \frac{\pi(1-p)}{\pi(1-p) + (1-\pi)p}$$

From the law of large numbers, a firm hiring N_h high-signal workers and N_l low-signal workers receives effective labor L equal to

$$N = \mathbb{P}(x(i) = S|s(i) = h)N_h + \mathbb{P}(x(i) = S|s(i) = l)N_l.$$

Therefore, firms choose N_h and N_l to solve

$$\max_{N_l, N_h} A \left[\frac{\pi p}{\pi p + (1-\pi)(1-p)} N_h + \frac{\pi(1-p)}{\pi(1-p) + (1-\pi)p} N_l \right] - w_h N_h - w_l N_l$$

The necessary and sufficient conditions for profit maximization are

$$w_h(\pi) = \frac{\pi p}{\pi p + (1-\pi)(1-p)} A \tag{2}$$

$$w_l(\pi) = \frac{\pi(1-p)}{\pi(1-p) + (1-\pi)p} A. \tag{3}$$

2.4 Equilibrium

Definition 1 *A Perfect Bayesian Nash Equilibrium is*

- *Household value and policy functions $V(a;\pi,\tau)$ and $e(a;\pi,\tau)$*
- *Wages $w_h(\pi)$ and $w_l(\pi)$*
- *Beliefs π^* about the skilled fraction of the population*

such that

1. ***Household Optimization:*** *The household value and policy functions solve the household's problem, (1).*
2. ***Firm Optimization:*** *Wages satisfy (2) – (3).*
3. ***Consistency of Beliefs:*** $\pi^* = \int e(a;\pi^*,\tau)d\Omega(a)$.

3 Theoretical Results

This section proves the existence of a Perfect Bayesian Nash Equilibrium (PBNE), analyzes its efficiency, and looks at the effects of estate taxes. First, make the following assumptions:

1. $u(c)$ is strictly increasing, continuously differentiable, and strictly concave; $v(e)$ is strictly increasing, twice continuously differentiable, and strictly convex.

2. $v(0) = 0$, $v'(0) = 0$ and $\lim_{e \to 1} v'(e) = +\infty$.

First, I establish some properties of household effort choice.

Lemma 1 *Effort $e(a; \pi, \tau)$ is a continuous, single-valued function in a and π.*

Lemma 2 *Effort $e(a; \pi, \tau)$ is interior and satisfies the first-order condition*

$$v'(e(a; \pi, \tau)) = (2p - 1)[u(w_h(\pi) + a(1 - \tau)) - u(w_l(\pi) + a(1 - \tau))]$$

The proofs of both lemmas are in the appendix.

3.1 Existence

Now I establish sufficient conditions for the existence of a PBNE with a non-trivial fraction $\pi^* > 0$ of skilled workers.

Theorem 1 (Existence of Perfect Bayesian Nash Equilibrium) *Given $\tau \in [0, 1]$, a non-trivial ($\pi^* > 0$) PBNE exists if*

$$\mathbb{E}_\Omega[u'(a(1 - \tau))] > \frac{p(1 - p)v''(0)}{(2p - 1)^2 A}$$

Proof. Define $\phi(\pi, \tau) = \int e(a; \pi, \tau) d\Omega(a)$ and $g(\pi, \tau) = \phi(\pi, \tau) - \pi$. A non-trivial equilibrium is $\pi^* \in (0, 1]$ such that $\phi(\pi^*) = \pi^*$, i.e. $g(\pi^*, \tau) = 0$.

Note that $w_h(0) = w_l(0)$ and $w_h(1) = w_l(1)$. Therefore, workers do not exert any effort if $\pi = 0$ or $\pi = 1$, i.e. $e(a; 0, \tau) = e(a; 1, \tau) = 0 \ \forall a \in [0, \overline{a}]$ and $\tau \in [0, 1]$. Therefore, $\phi(0) = \phi(1) = 0$, implying that $g(0) = 0$ and $g(1) = -1 < 0$.

To show existence of $\pi^* > 0$, I first show that $\exists \pi \in (0, 1)$ such that $g(\pi, \tau) > 0$. Then, by the Intermediate Value Theorem, $\exists \pi^* \in (\pi, 1]$ such that $g(\pi^*, \tau) = 0$. To show the first claim, it suffices to show that $\phi(\pi,\tau)$ is differentiable in π at the origin and $\phi_\pi(0,\tau) > 1$.

Define $f(e,\pi,\tau;a) = v'(e) - (2p-1)[u(w_h(\pi) + a(1-\tau)) - u(w_l(\pi) + a(1-\tau))]$. The first-order condition for $e(a;\pi,\tau)$ becomes $f(e,\pi,\tau;a) = 0$ $\forall a \in [0,\overline{a}]$.

Because $v \in C^2$ and $u \in C^1$, $f \in C^1$ with $f(0,0,\tau;a) = 0$. Thus, given a and τ, the Implicit Function Theorem states that there exists an open neighborhood $\mathcal{N}_\varepsilon(0)$ about $\pi = 0$ for some $\varepsilon > 0$ such that $e(a;\pi,\tau) \in C^1$ $\forall \pi \in \mathcal{N}_\varepsilon(0) \bigcap [0,1]$. Furthermore,

$$\left.\frac{\partial e}{\partial \pi}\right|_{\pi=0} = -\left.\frac{f_\pi}{f_e}\right|_{\pi=0} = \frac{1}{v''(\underbrace{e(a;0,\tau)}_{0})}(2p-1)u'(a(1-\tau))\left.\left(\frac{\partial w_h}{\partial \pi} - \frac{\partial w_l}{\partial \pi}\right)\right|_{\pi=0}.$$

Differentiating $w_h(\pi) - w_l(\pi) = \dfrac{\pi p}{\pi p + (1-\pi)(1-p)}A - \dfrac{\pi(1-p)}{\pi(1-p) + (1-\pi)p}A$ at $\pi = 0$ gives

$$\left.\left(\frac{\partial w_h}{\partial \pi} - \frac{\partial w_l}{\partial \pi}\right)\right|_{\pi=0} = \frac{(2p-1)A}{p(1-p)}.$$

Therefore, $\left.\dfrac{\partial e}{\partial \pi}\right|_{\pi=0} = \dfrac{(2p-1)^2 A u'(a(1-\tau))}{p(1-p)v''(0)}$.

The above expression holds for all a, and because $e(a;\pi,\tau)$ is a bounded, continuous (and thus measurable) function, it follows that ϕ is differentiable at the origin and

$$\phi_\pi(0,\tau) = \left.\frac{\partial}{\partial \pi}\right|_{\pi=0}\int e(a;\pi,\tau)d\Omega(a) = \int \left.\frac{\partial e(a;\pi,\tau)}{\partial \pi}\right|_{\pi=0} d\Omega(a)$$
$$= \int \frac{(2p-1)^2 A u'(a(1-\tau))}{p(1-p)v''(0)}d\Omega(a) = \mathbb{E}_\Omega\left[\frac{(2p-1)^2 A u'(a(1-\tau))}{p(1-p)v''(0)}\right].$$

Therefore, a non-trivial PBNE exists if

$$\phi_\pi(0,\tau) > 1 \Leftrightarrow \mathbb{E}_\Omega[u'(a(1-\tau))] > \frac{p(1-p)v''(0)}{(2p-1)^2 A}.$$

■

3.2 Efficiency and Welfare

The marginal cost of skill acquisition effort is $v'(e)$, and the marginal private benefit is

$$I(\pi, a, \tau) = (2p - 1)[u(w_h(\pi) + a(1 - \tau)) - u(w_l(\pi) + a(1 - \tau))].$$

By increasing effort, workers are more likely to acquire skills, and therefore more likely to send a high signal to the labor market and receive a higher wage. Wages are higher for high-signal workers precisely because firms know that such workers have a greater probability of being skilled. Note that $I(\pi, a, \tau)$ also depends on the population fraction π of skilled workers, with $I(\pi, a, \tau) = 0$ when $\pi = 0$ and also when $\pi = 1$. Therefore, for low π, an increase in π increases the incentive to exert effort, which makes effort and π complementary. However, for high π, further increases in π decrease the incentive to exert effort.

In either case, π is a public good because each worker's effort choice contributes to π, and all workers benefit from higher π. Workers, however, do not consider the impact of their effort on π and, therefore, under-invest in effort. In short, the imprecise signaling of skills gives rise to an informational externality that causes workers to engage in informational free-riding. Theorem 2 formalizes this inefficiency result.

Theorem 2 (Inefficiency of No-Tax Equilibrium) *The PBNE with $\tau = 0$ is inefficient with too few skilled workers, i.e. $\pi^* < \pi^{opt}$.*

Proof. Recall that, given π, a worker with assets a chooses effort $e(a; \pi, 0)$ such that

$$v'(e(a; \pi, 0)) = I(\pi, a, 0) = (2p - 1)[u(w_h(\pi) + a) - u(w_l(\pi) + a)]$$

Unlike workers, a utilitarian social planner problem internalizes the effect of each worker's effort choice on aggregate skilled labor π, solving

$$W = \max_{\pi, e(a)} \int \{[e(a)p + (1-e(a))(1-p)]u(w_h(\pi)+a)$$
$$+[e(a)(1-p) + (1-e(a))p]u(w_l(\pi)+a) - v(e)\}\, d\Omega(a)$$
subject to
$$\pi = \int e(a) d\Omega(a)$$

The social planner's choice $e(a)$ satisfies the first-order condition

$$v'(e(a)) = \underbrace{(2p-1)[u(w_h(\pi^{opt})+a) - u(w_l(\pi^{opt})+a)]}_{=I(\pi^{opt},a,0)=v'(e(a;\pi^{opt},0))} + f_\Omega(a)\,\{[e(a)p + (1-e(a))(1-p)]$$
$$\times u'(w_h(\pi^{opt})+a)\frac{\partial w_h}{\partial \pi} + [e(a)(1-p) + (1-e(a))p]u'(w_l(\pi^{opt})+a)\frac{\partial w_l}{\partial \pi}\Bigg\}$$

where $\pi^{opt} = \int e(a)d\Omega(a)$.

The first term corresponds to the first-order condition of the worker's effort choice when $\pi = \pi^{opt}$. Therefore, $v'(e(a)) > v'(e(a;\pi^{opt},0))$ because the derivatives $u' > 0$, $\frac{\partial w_h}{\partial \pi} > 0$, and $\frac{\partial w_l}{\partial \pi} > 0$. From the strict convexity of the effort cost function, the condition $v'(e(a)) > v'(e(a;\pi^{opt},0))$ implies that $e(a) > e(a;\pi^{opt},0)$.

Therefore, $\pi^{opt} = \int e(a)d\Omega(a) > \int e(a;\pi^{opt},0)d\Omega(a)$, implying that $\phi(\pi^{opt},0) < \pi^{opt}$ and $g(\pi^{opt},0) < 0$, where ϕ and g are as defined in the proof of equilibrium existence. Because $g(\pi,0) \geq 0\ \forall \pi \in [0,\pi^*]$, it must be that $\pi^{opt} > \pi^*$. Therefore, the optimal skilled fraction of the population is greater than the equilibrium skilled fraction. ■

3.3 Inheritance, Effort, and the Estate Tax

With risk-averse workers, inheritances exacerbate the effort distortion from the information externality and reduce the incentive to exert effort. Lemma 3 establishes this result formally.

Lemma 3 *Effort $e(a;\pi,\tau)$ is decreasing in assets a and increasing in the estate tax rate τ.*

The proof is in the appendix. By decreasing inheritances, an estate tax mitigates the negative effect of wealth on effort and increases the proportion π of skilled workers in the population. Because of the information externality, an increase in π raises wages for both high-signal and low-signal workers. Lemma 4 formalizes this result and is proved in the appendix.

Definition 2 *Let the (non-trivial) equilibrium skilled proportion of the population as a function of the tax rate be given implicitly by $\pi(\tau) = \phi(\pi(\tau), \tau)$, where $\phi(\pi, \tau) = \int e(a; \pi, \tau) d\Omega(a)$.*

Lemma 4 *If $\phi_\pi(0, 0) > 1$ (which is equivalent to substituting $\tau = 0$ into the condition in theorem 1), then the equilibrium skilled proportion of the population increases in the estate tax rate, i.e. $\pi(\tau') > \pi(\tau)$ for all $\tau' > \tau$.*

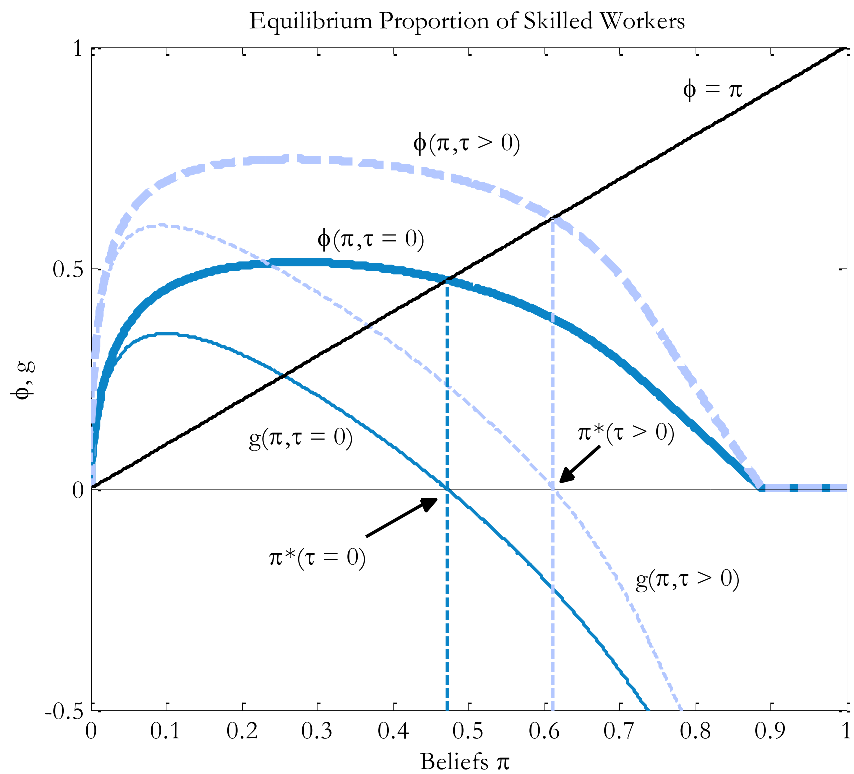

Figure 1: The equilibrium proportion of skilled workers for $\tau = 0$ and $\tau > 0$. Equilibrium $\pi^* \equiv \pi(\tau)$ occurs where $\pi(\tau) = \phi(\pi(\tau), \tau)$, or equivalently, $g(\pi(\tau), \tau) \equiv \phi(\pi(\tau), \tau) - \pi(\tau) = 0$.

Theorem 3 below demonstrates that an increase in the estate tax leads to welfare gains for low-asset households, *even when all the revenue is used for wasteful government spending.* By increasing the equilibrium fraction of skilled workers, the estate tax shrinks the information externality and causes wages to increase for all workers. For low-inheritance households, the benefit of higher wages more than compensates for the loss in initial wealth, making the estate tax a progressive policy intervention.

Definition 3 *Let $V(a; \pi(\tau), \tau)$ be equilibrium welfare for a worker with a taxed at rate τ.*

Theorem 3 *If $\phi_\pi(0,0) > 1$, then $\forall \tau \in [0,1)$ $\exists a(\tau) \in (0, \overline{a}]$ such that $\frac{dV(a;\pi(\tau),\tau)}{d\tau} > 0$ $\forall a < a(\tau)$. In words, there is a positive measure of workers with inheritance $a < a(\tau)$ for whom a small increase in the estate tax is welfare improving.*

Proof. From previous results that $e(a;\pi,\tau) \in C^1$ and $\pi(\tau) \in C^1$, $V(a;\pi(\tau),\tau)$ is continuously differentiable. Now consider a worker with no inheritance, namely, $a = 0$. Totally differentiating $V(0;\pi(\tau),\tau)$ gives

$$\frac{dV(0;\pi(\tau),\tau)}{d\tau} = \frac{\partial V(0;\pi(\tau),\tau)}{\partial \pi}\pi'(\tau) + \frac{\partial V(0;\pi(\tau),\tau)}{\partial \tau}.$$

For this worker with no inheritance,

$$V(0;\pi(\tau),\tau) = \max_{e\in[0,1]}[ep + (1-e)(1-p)]u(w_h(\pi(\tau))) + [e(1-p) + (1-e)p]u(w_l(\pi(\tau))) - v(e)$$

The envelope theorem and the fact that τ does not directly appear in the right-hand side of the above expression imply that

$$\begin{aligned}
\frac{\partial V(0;\pi(\tau),\tau)}{\partial \tau} &= 0 \\
\frac{\partial V(0;\pi(\tau),\tau)}{\partial \pi} &= [e(0;\pi(\tau),\tau)p + (1-e(0;\pi(\tau),\tau))(1-p)]u'(w_h(\pi(\tau)))\frac{\partial w_h}{\partial \pi} \\
&\quad + [e(0;\pi(\tau),\tau)(1-p) + (1-e(0;\pi(\tau),\tau))p]u'(w_l(\pi(\tau)))\frac{\partial w_l}{\partial \pi}
\end{aligned}$$

Therefore, $\frac{dV(0;\pi(\tau),\tau)}{d\tau} > 0$ because $u' > 0$, $\frac{\partial w_h}{\partial \pi} > 0$, $\frac{\partial w_l}{\partial \pi} > 0$, and $\pi'(\tau) > 0$.

Now suppose that $\exists \widehat{a} \in (0,\overline{a}]$ such that $\frac{dV(a;\pi(\tau),\tau)}{d\tau} \leq 0$ (otherwise the desired result trivially holds). By the Intermediate Value Theorem, $\exists a(\tau) \in (0,\widehat{a}]$ such that $\frac{dV(a(\tau);\pi(\tau),\tau)}{d\tau} = 0$, and by the continuity of $\frac{dV(a;\pi(\tau),\tau)}{d\tau}$, the result is $\frac{dV(a;\pi(\tau),\tau)}{d\tau} > 0$ $\forall a < a(\tau)$. ■

This theorem neither establishes whether a majority of households benefit from a higher estate tax nor gives an indication about the optimal tax rate. Furthermore, the static model ignores employer learning and the effects of estate taxation on capital accumulation and bequests. I investigate these issues in the remainder of the paper.

4 The Dynamic Model

This section builds upon the one period model and extends it to a dynamic environment that includes Bayesian learning of worker skills and capital accumulation. In the subsequent quantitative section, I use a parametrized version of the model that also includes Lucas (1988) human capital externalities to evaluate the effects of estate taxation and to determine the optimal tax rate.

4.1 Households

Households face an infinite horizon but are subject to constant mortality risk with survival probability φ each period. When a household with assets a dies, a new household (youth) replaces it, inherits a, and pays the estate tax τ. Youth make a *one-time* effort choice e to stochastically acquire skills,

$$x = \begin{cases} S & \text{with probability } e \\ U & \text{with probability } 1 - e. \end{cases}$$

Households retain their skill status until they die. All workers inelastically supply one unit of time to the labor market, but effective labor for unskilled workers is less than that of skilled workers: $n_U < n_S \equiv 1$. Effort disutility is $v(e)$, period consumption utility is $u(c)$, and households discount at the rate β and exhibit perfect altruism toward their descendants.

Each period, workers send an imprecise signal $s \in \{l, h\}$ to their employer about their skill level. The accuracy of the signal satisfies

$$p = \mathbb{P}(s = h | x = S) = \mathbb{P}(s = l | x = U) > \frac{1}{2},$$

and the signals are iid across households and time.

These stochastic labor market signals affect the wage that a worker receives and, therefore, act as a source of idiosyncratic risk. Financial markets are incomplete, with workers only able to self-insure by accumulating assets (capital). Households lend capital to firms and earn rate of return r.

4.2 Firms

Firms operate a constant returns to scale technology using capital K and effective labor N,

$$Y = AF(K, N).$$

Firms rely on the signals to infer each worker's skill level. However, to allow for learning over time, employment relationships are long-lasting with separation probability λ.

4.3 Informational Environment

To prevent asymmetric information between firms, I assume that the history of signals each worker receives *while employed at their current employer* is publicly observable to all firms. As a result, firms engage in Bertrand competition for workers based on signal histories $s^\tau = \{s_{-\tau+1}, \ldots, s_{-1}, s\}$, where τ is the length of the history. However, if a worker-firm relationship dissolves, new employers can only observe the worker's initial signal and not the histories from previous employment matches.[1] As in the static model, the share π of skilled workers in the economy plays a key role in belief formation and wage determination.

4.3.1 Bayesian Learning

Firms enter each period with prior beliefs about the skill level of each worker. For workers with no signal history (because of a recent job separation or because they began the period as youth), $\mu^0 \equiv \mathbb{P}(x = S|\emptyset) = \pi$, where π is the fraction of skilled workers in the population.

The posterior belief for such workers after receiving their first signal is

$$\mu^1 \equiv \mathbb{P}(x = S|s_1) = \frac{\pi p_{s_1}}{\pi p_{s_1} + (1-\pi)(1-p_{s_1})}, \text{ where } p_{s_1} \equiv p\mathbf{1}_{[s_1=h]} + (1-p)\mathbf{1}_{[s_1=l]}.$$

In general, define $\mu^\tau \equiv \mathbb{P}(x = S|s^\tau)$. The next theorem shows that, rather than keep track of the entire history of signals s^τ, firms and workers can use μ^τ as a sufficient statistic for updating beliefs about skills and compensation.

[1] The parameter λ is calibrated to reflect only involuntary separations. Implicitly, I assume workers keep their signal histories in job-to-job transitions.

Theorem 4 (Bayesian Learning of Worker Skills) *Given prior beliefs $\mu^{\tau-1}$ and signal s, the posterior likelihood μ^{τ} that a worker is skilled is given by*

$$\mu^{\tau} = \frac{\mu^{\tau-1}p_s}{\mu^{\tau-1}p_s + (1-\mu^{\tau-1})(1-p_s)}$$

where $p_s \equiv p\mathbf{1}_{[s=h]} + (1-p)\mathbf{1}_{[s=l]}$.

The proof of theorem 4 is in the appendix.

4.4 Government Policy

The government taxes inherited wealth at the rate τ_a. I consider four possible uses of the revenue: wasteful government consumption, lump-sum rebates, wage subsidies for all youth, or wage subsidies only for youth who receive a high signal. The latter is analogous to a higher education subsidy or tax break. Let T_s denote the transfer to youth with signal s.

4.5 Decision Problems

Each period is divided into two subperiods. In subperiod 1, youth receive their inheritance a, pay the estate tax, and exert skill effort e. Workers enter the period having just learned whether they experienced a job separation and erasure of their signal history or whether they remain employed at their current firm.

In subperiod 2, youth learn the outcome of their stochastic skill acquisition and become workers. Furthermore, *all* workers send a new signal to their employers and make consumption c and savings decisions, $a' \geq 0$. Now I describe the value functions in steady state.

4.5.1 Household Problem

Workers enter *subperiod 2* with assets a and an updated μ that summarizes their signal history. They have value function

$$V_x(a,\mu) = \max_{c,a'\geq 0} u(c) + \beta\left[(1-\varphi)Y(a') + \varphi\left\{(1-\lambda)W_x(a',\mu) + \lambda W_x(a',\pi)\right\}\right]$$

$$\text{subject to} \tag{4}$$

$$c + a' = w\left[(1-\mu)n_U + \mu\right] + (1+r)a$$

The first continuation term is youth utility and reflects the worker's bequest motive in the event of death. The last two terms give future utility conditional on job retention or separation. In the latter case, the worker's signal history is erased and replaced by the fraction π of skilled workers in the entire population, which is common knowledge.

The continuation term $W_x(a',\mu)$ represents beginning-of-period utility and satisfies

$$\begin{aligned} W_S(a',\mu) &= pV_S(a',\mu'_{\mu,h}) + (1-p)V_S(a',\mu'_{\mu,l}) \\ W_U(a',\mu) &= pV_U(a',\mu'_{\mu,l}) + (1-p)V_U(a',\mu'_{\mu,h}) \end{aligned} \tag{5}$$

where beliefs about the worker's skill level are updated as follows:

$$\mu'_{\mu,s} \equiv \mu'(\mu,s) = \frac{\mu p_s}{\mu p_s + (1-\mu)(1-p_s)}, \text{ with } p_s = p\mathbf{1}_{[s=h]} + (1-p)\mathbf{1}_{[s=l]}. \tag{6}$$

Recall that beliefs are updated in between subperiods 1 and 2, *after* youth have made their effort choice and learned the outcomes of their stochastic skill acquisition. Corresponding to this choice, youth have value function

$$Y(a) = \max_{e\in[0,1]} -v(e) + \left[\begin{array}{l} epV_S(a(1-\tau_a)+T_h,\mu'_{\pi,h}) + e(1-p)V_S(a(1-\tau_a)+T_l,\mu'_{\pi,l}) \\ (1-e)pV_U(a(1-\tau_a)+T_l,\mu'_{\pi,l}) + (1-e)(1-p)V_U(a(1-\tau_a)+T_h,\mu'_{\pi,h}) \end{array}\right] \tag{7}$$

Note that the skilled population share π acts as the firm's prior belief for youth just entering the labor force and for recently separated workers. The firm then forms posterior beliefs $\mu'_{\pi,s}$ based on observing the worker's first signal s. The population share π holds no informational content for workers and firms who remain matched from the previous period.

4.5.2 Firm's Problem

As described in section 4.3, *all* firms observe the history of signals received by each worker *while at their current employer*. Incumbent firms have no informational advantage.

Firms choose capital and the amount of workers with expected skill-level μ to solve

$$\max_{K,\{N(\mu)\}} AF\left(K, \int \left[(1-\mu)n_U + \mu\right] N(d\mu)\right) - \int w(\mu)N(d\mu) - (r+\delta)K$$

The necessary and sufficient conditions for profit maximization are

$$r + \delta = AF_K \tag{8}$$

$$w(\mu) = \left[(1-\mu)n_U + \mu\right] AF_N \tag{9}$$

4.6 Equilibrium

In words, a *Stationary Recursive Perfect Bayesian Nash Equilibrium* is a collection of value and policy functions, $V_x(a,\mu)$, $W_x(a',\mu)$, $Y(a)$, $c_x(a,\mu)$, $a'_x(a,\mu)$, and $e(a)$, market prices r and $w(\mu)$, aggregate quantities K and N, beliefs π about the population share of skilled workers, and measures $\Lambda_Y(a)$ and $\Lambda_W(a,\mu)$ such that households maximize utility, firms maximize profits, markets clear, beliefs are consistent with reality, and the measures are invariant with respect to the Markov process induced by stochastic skill acquisition, signaling, separations, and all relevant policy functions.

5 Quantitative Results

This section uses a parametrized version of the dynamic model to analyze the macroeconomic effects of information externalities. After describing the parametrization, I demonstrate how the degree of labor market signal inaccuracy affects equilibrium prices and allocations. From there, I assess the ability of estate taxation to mitigate the effects of the information externality. Lastly, I determine the optimal estate tax under a variety of scenarios, including one that supplements the baseline model with Lucas (1988) human capital externalities.

Table 1: Parameter Summary

Description	Parameter	Value	Target	Model	Target Description
		Independent Parameters			
Risk Aversion	σ	2			Standard
Effort Cost Exponent	η	1.5			Convex Disutility
Survival Rate	φ	0.975			40 Years of Working Life
Unskilled Productivity	n_U	0.5			$n_S/n_U = 2$ (100% College Premium)
Separation Probability	λ	0.133			Annual Involuntary Separation Rate*
Capital Share	α	0.36			36% Capital Share
Depreciation Rate	δ	0.1			10% Annual Depreciation
Signal Accuracy**	p	0.7			
		Jointly Determined Parameters			
TFP	A	0.8851	1	1.0	Annual Output (Normalization)
Discount Rate	β	0.9773	3	3.0	Capital-Output Ratio
Effort Cost Coefficient	ψ	0.5185	30%	30.3%	% Skilled (U.S. College Graduate %)

*Annualized involuntary separation rate adapted from figure 8 in Shimer (2005).

**Robustness checks are performed by also using $p = 0.6$ and $p = 0.8$.

5.1 Parametrization

Some parameters of the model are set using a priori information, while the remaining parameters are determined jointly within the model. The baseline estate tax is $\tau_a = 0$.

Households Consumption felicity and effort disutility are $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $v(e) = \psi e^{\eta}$, respectively, with standard risk aversion of $\sigma = 2$ and $\eta = 1.5$ set to ensure convexity of effort costs. The parameter ψ is determined jointly to equate the equilibrium skilled population share with the 30% share of college graduates in the U.S. The survival probability is $\varphi = 0.975$ to yield an average working life of 40 years, and unskilled worker productivity is $n_U = 0.5$ (implying $n_S/n_U = 2$) to match the 100% college earnings premium. The discount factor β is set jointly to match the capital-output ratio of 3.

Firms Production is Cobb-Douglas, $Y = AK^{\alpha}L^{1-\alpha}$, with a 36% capital share and 10% depreciation rate. I determine A internally to normalize baseline output to 1.

Informational Environment For the signal accuracy, I use a baseline value of $p = 0.7$ but later consider $p = 0.6$ and $p = 0.8$ for robustness.

5.2 Information Externalities in the Dynamic Economy

Just as in the static model, imperfect signaling creates an information externality in the labor market. Specifically, firms augment the private signals with their knowledge of the skilled population share to form beliefs over the skill level of each individual worker. However, workers do not internalize the public good component of their skill acquisition when making effort choices. The dynamic economy also augments the static model by introducing learning and endogenous saving from intergenerational altruism and precautionary behavior.

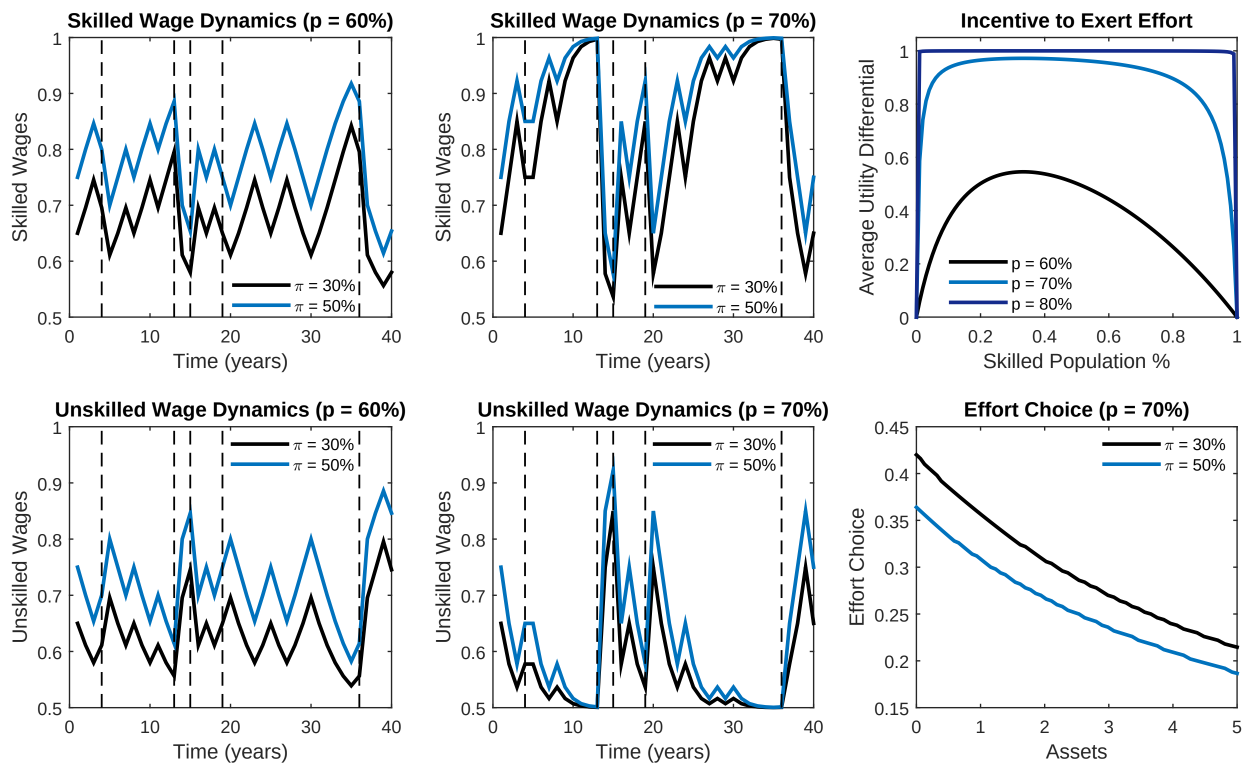

Figure 2: Wage dynamics, the incentive to exert effort, and youth effort choice for different signal accuracies and skilled population shares. The vertical dashed lines depict separations.

Even with learning, firms' knowledge of the skilled population π has a *persistent* effect on the wages of individual workers. However, as one might expect, higher signal accuracy p increases the speed of learning and reduces the magnitude of the information externality. Figure 2 shows simulated wage dynamics for skilled and unskilled workers for different values of p and π. For the sake of comparison, I assume that whenever a skilled worker draws an accurate (inaccurate) signal, the unskilled worker also draws an accurate (inaccurate) signal. For example, if the skilled worker draws $s_t = h$ in period t, the unskilled worker draws $s_t = l$.

The first column shows that learning is quite slow when $p = 60\%$, with wages exhibiting only modest convergence between job separations toward each worker's true productivity. Furthermore, the effect of the skilled share π on wages is large and persistent. By contrast, learning is more noticeable when $p = 70\%$, as shown in the second column of panels. In between job separations, wages of skilled workers converge to 1, while the wages of unskilled workers converge to 0.5. Notably, the skilled share π still has a persistent, though attenuated, effect on wages when $p = 70\%$.

The top right panel gives a visual representation of the private incentive to exert effort as a function of p and π. Specifically, given simulated sequences of wages $\{w_t^S(p,\pi)\}_{t=0}^T$ and $\{w_t^U(p,\pi)\}_{t=0}^T$, this panel plots the discounted average utility differential,[2]

$$I(p,\pi) = (1-\beta)\sum_{t=0}^{T}\beta^t[u(w_t^S(p,\pi)) - u(w_t^U(p,\pi))].$$

Regardless of p, there is no incentive to exert any effort when $\pi = 0\%$ or $\pi = 100\%$ because firms have degenerate prior beliefs. For interior π, firms actively learn by incorporating knowledge of π with the observed signals from each worker. When $p = 60\%$, the skilled worker share π has substantial informational content and greatly influences the incentive to invest. Although the maximum average wage differential between skilled and unskilled workers is maximized at $\pi = 50\%$, risk aversion makes it such that the maximum average *utility* differential occurs between $\pi = 30\%$ and $\pi = 40\%$. When $p = 70\%$ or $p = 80\%$, signals are much more informative to firms. As a result, the private incentive to exert effort to acquire skills is higher *and less responsive* to the skilled worker share π.

Lastly, the bottom right panel shows effort choice as a function of assets. Assets decrease the incentive to exert effort for two reasons. First, effort declines due to a standard wealth effect from the diminishing marginal utility of consumption. Second, increased asset holdings reduce the *precautionary* incentive to exert effort. Because signals, and therefore wages, are stochastic and uninsurable, workers face significant consumption risk. Without any assets, youth have a strong incentive to acquire skills to reduce fluctuations in marginal utility. By contrast, wealthier households can simply use their assets to smooth consumption.

[2] The true incentive to exert effort is given in equation (7).

5.2.1 The Equilibrium Response to Higher Signal Accuracy

Table 2 gives further insight into the effects of signal accuracy by showing the steady state response to an increase in p from 70% to 80%. The enhanced incentive to acquire skills causes the fraction of skilled workers to increase from 30.3% to 53.1%. In response, starting wages increase for *all* workers. For workers who initially receive a high signal, firms assign a stronger posterior probability to the worker being skilled both because of the higher signal accuracy and because of the greater share of skilled workers π. In isolation, higher signal accuracy reduces the starting wages of low-signal workers, but this decrease is more than offset by the effect of the higher population share π on wages. Overall, starting wages rise from 0.74 to 0.9 for high-signal workers and from 0.57 to 0.61 for low-signal workers.

Table 2: Equilibrium Response to Higher Signal Accuracy

	$p = 70\%$	$p = 80\%^*$
Skilled Worker Share	30.3%	53.1%
Risk-Free Rate	1.97%	1.76%
Aggregate Output	1.00	1.19
Capital-Output Ratio	3.00	3.06
High-Signal Starting Wage	0.74	0.90
Low-Signal Starting Wage	0.57	0.61

*All other parameters are held the same.

These higher wages allow workers to increase their capital accumulation in absolute terms. However, higher precautionary saving is needed to explain the increase in capital *as a fraction of output* from 3 to 3.06. In general, higher signal accuracy has two opposing effects on the precautionary savings motive. On the one hand, higher signal accuracy reduces the frequency of signal fluctuations. On the other hand, wages are more responsive to signal fluctuations, as shown in the comparison of the $p = 60\%$ and $p = 70\%$ panels in figure 2. In other words, wages fluctuate between a wider band of values when p is higher. Overall, this latter effect dominates here and causes an increase in precautionary saving. That said, an increase in p all the way to 100% would completely eliminate idiosyncratic risk for workers. Workers would always receive accurate signals, and those signals would completely reveal the worker's type—leaving no role for the skilled worker share π in wage determination.

5.3 The Effects of Estate Taxation

This section analyzes the steady state equilibrium response to an imposition of a 50% estate tax relative to the no-tax baseline. I consider several different scenarios for how the government uses the tax revenue, and I also analyze the role of signal accuracy, altruism, and the impact of adding Lucas (1988) human capital externalities to the model. The next section computes the optimal estate tax in a variety of circumstances.

5.3.1 Estate Taxation under Different Transfer Schemes

I consider four possible uses of the tax revenue: wasteful government consumption, lump-sum rebates to youth, wage subsidies to youth, and wage subsidies only to youth who receive a high signal upon entering the labor market.[3] The government runs a balanced budget.

Table 3: Effects of a 50% Estate Tax: Different Transfer Schemes

	No Transfer	Lump-Sum	Subsidy	$\text{Subsidy}_{\text{targeted}}$
Aggregates				
Skilled Workers	35.8%	27.7%	29.1%	40.2%
Risk-Free Rate	3.33%	3.29%	3.27%	3.09%
Output	0.98	0.92	0.93	1.02
Capital/Output	2.70	2.71	2.71	2.75
Δ*Full Info* MP_N	−5.88%	−5.73%	−5.66%	−4.90%
Δ*High-Signal Wage*	−1.96%	−7.70%	−6.58%	1.89%
Δ*Low-Signal Wage*	−2.93%	−7.03%	−6.28%	0.58%
Welfare Change*				
All	−4.32%	−3.68%	−2.94%	2.39%
Youth	−8.32%	−2.93%	−2.41%	0.60%

*Welfare is measured in consumption-equivalent units.

The choice of how to allocate the revenues from an estate tax has first order equilibrium impacts, as shown in table 3. If the government simply spends the revenues on wasteful government consumption, the skilled share rises from 30.3% to 35.8%. The negative wealth effect from the reduction in inherited wealth induces youth to exert greater effort, and the rising share of skilled workers π reinforces the stronger incentive to acquire skills. At the

[3] For example, subsidies that are targeted to college graduates.

same time, the estate tax has a negative effect on saving and causes the capital to output ratio to fall from 3 to 2.7. The increase in skilled workers and reduction in capital have opposing effects on output. Overall, output decreases by 2% and starting wages decline for both types of workers. Lower wages and wealth cause consumption-equivalent welfare to fall by 4.32% for workers and by 8.32% for youth.

If the government rebates the tax revenue lump sum to new workers, the policy still imposes welfare losses and completely fails to achieve its objective of increasing the share of skilled workers. In fact, the skilled worker population actually *falls* from 30.3% to 27.7%. This decline arises primarily because of the convexity of the effort policy function. Specifically, effort as a function of wealth falls fastest for low levels of assets. Therefore, even ignoring the endogenous response of saving, an estate tax with lump sum rebates acts as a mean preserving compression of the asset distribution. The increased effort from impoverishing wealthy youth is more than offset by the reduction in effort from enriching poor youth. The drop in skilled workers, when combined with the reduction in capital accumulation, causes output to fall by 8% when the revenue is rebated lump sum—far more than when the revenues were simply spent on waste.

The universal wage subsidy for new workers improves modestly on the lump sum rebates but still has qualitatively the same effect. Only when the government directs all revenues to targeted subsidies for new workers who initially receive a high signal does the estate tax show significant promise. In this last case, the skilled share increases dramatically from 30.3% to 40.2%, and output actually *increases* by 2% relative to its initial steady state. Starting wages also increase modestly—despite a drop in the marginal product of labor—and worker welfare improves by 2.39%. Youth experience more modest gains of 0.60% of lifetime consumption.

5.3.2 Estate Taxation under Different Signal Accuracies

Given the importance of signal accuracy for the magnitude of the information externality, I now conduct some robustness checks by analyzing the effects of a 50% estate tax when $p \in \{60\%, 80\%\}$. In both cases, I recalibrate the model to match a 30% skilled worker share, a capital to output ratio of 3, and normalized annual output of 1, as in table 1. From here forward, I assume the government distributes the tax revenues using targeted subsidies.

Table 4: Effects of a 50% Estate Tax: Different Signal Accuracies*

	$p = 60\%$	$p_{\text{baseline}} = 70\%$	$p = 80\%$
Aggregates			
Skilled Workers	42.9%	40.2%	40.8%
Risk-Free Rate	3.15%	3.09%	3.08%
Output	1.04	1.02	1.03
Capital/Output	2.74	2.75	2.75
Δ*Full Info* MP_N	−5.05%	−4.90%	−4.75%
Δ*High-Signal Wage*	4.53%	1.89%	1.20%
Δ*Low-Signal Wage*	3.71%	0.58%	−0.41%
Welfare Change			
All	6.44%	2.39%	0.99%
Youth	5.41%	0.60%	−1.82%

*The model is recalibrated each time to match $\pi = 30\%$, $Y = 1$, and $K/Y = 3$ when $\tau_a = 0$.

Raising signal accuracy to 80% does not significantly alter the response of the skilled worker share, which rises from 30% to 40.8% instead of 40.2%. Furthermore, output increases by 3% instead of 2%. However, perhaps unsurprisingly, the welfare gains for workers decrease significantly from 2.39% to 0.99%. Furthermore, when $p = 80\%$, the tax-induced drop in low-signal wages combined with the loss of wealth and increased effort disutility cause youth to actually experience welfare *losses* of 1.82%.

However, lower signal accuracy of $p = 60\%$ magnifies the equilibrium response and welfare effects of the tax, with the skilled population increasing by 12.9% instead of 10.2%. Recall from the top right panel of figure 2 that the incentive to acquire skills is more sensitive to the skilled population share π when labor market signaling is less accurate. Therefore, as effort increases in response to diminished inheritances, the resulting rise in π reinforces the incentive to exert effort and magnifies the increase in skilled workers.

Lastly, note that the gap between the full information marginal product of labor and starting wages is decreasing in signal accuracy. In all three scenarios, the marginal product of labor falls substantially because of higher effective labor and lower capital. However, by reducing informational free-riding and increasing the skilled population π, estate taxation can in some circumstances reduce the information wedge by enough to increase wages.

Table 5: Effects of a 50% Estate Tax: Altruism and Human Capital Externalities*

	Baseline	No Altruism	Lucas I	Lucas II
Aggregates				
Skilled Workers	40.2%	35.3%	38.2%	36.3%
Risk-Free Rate	3.09%	2.00%	3.09%	3.09%
Output	1.02	1.04	1.05	1.07
Capital/Output	2.75	3.00	2.75	2.75
Δ*Full Info* MP_N	−4.90%	−0.04%	−1.39%	2.10%
Δ*High-Signal Wage*	1.89%	2.99%	3.07%	6.69%
Δ*Low-Signal Wage*	0.58%	4.04%	4.31%	5.59%
Welfare Change				
All	2.39%	4.05%	4.92%	7.45%
Youth	0.60%	1.07%	3.31%	6.01%

*The model is recalibrated each time to match $\pi = 30\%$, $Y = 1$, and $K/Y = 3$ when $\tau_a = 0$. Targeted subsidies are used when $\tau = 50\%$. Lucas I has $Y = \widetilde{A}\pi^{\phi}K^{\alpha}N^{1-\alpha}$ with $\phi = 0.1$. Lucas II sets $\phi = 0.25$.

5.3.3 The Role of Altruism and Human Capital Externalities

Altruism and Savings Until now, estate taxation has had two opposing effects on output. On the one hand, the shared of skilled workers rises, and thus, so does aggregate labor. On the other hand, the capital stock shrinks as a fraction of output because of a reduced incentive to save. When households anticipate that their children will face an estate tax, the bequest motive shrinks. To eliminate the effect of reduced capital accumulation, I now assume that households do not have any bequest motive at all. Instead, in the absence of perfect annuity markets, households only give bequests "accidentally" when they die with positive savings. This assumption implies the following change to workers' subperiod 2 value function:

$$V_x(a,\mu) = \max_{c,a'\geq 0} u(c) + \beta\varphi\left[(1-\lambda)W_x(a',\mu) + \lambda W_x(a',\pi)\right]$$

$$\text{subject to} \tag{10}$$

$$c + a' = w\left[(1-\mu)n_U + \mu\right] + (1+r)a.$$

With this formulation, the estate tax no longer distorts savings because it only impacts the youth value function $Y(a)$, which no longer appears in the continuation utility of workers.

Table 5 confirms this intuition by showing that a 50% estate tax results in no change to the capital to output ratio. However, precisely because capital remains elevated, the diminished wealth effect implies a smaller response of the skilled population share—a 5.3% increase in this case compared to the 10.2% surge in the baseline economy with altruistic bequests. Nevertheless, removing the savings distortion causes the estate tax to almost double its positive impact on output, wages, and welfare.

Human Capital Externalities Thus far, skill acquisition has had a public good component because of the information externality from imperfect labor market signaling. However, Lucas (1988) proposes an alternative mechanism to generate positive spillovers from skill acquisition. In his model, TFP depends on the endogenous stock of human capital, but workers do not internalize how their skill investments impact TFP. To analyze the interaction of this channel with the information externality emphasized in this paper, I modify the production function to allow TFP to depend on the share of skilled workers,

$$Y = \underbrace{\widetilde{A}\pi^{\phi}}_{A} K^{\alpha} L^{1-\alpha}$$

In the baseline model without Lucas externalities, $\phi = 0$ and $A \equiv \widetilde{A}$ is calibrated to normalize output to 1. Considerable disagreement exists in the literature regarding the magnitude of ϕ. For example, Iranzo and Peri (2009) argue that an increase of one college year per worker causes TFP to rise by 6%, and Gennaioli et al. (2013) claim that a 1% increase in the share of college graduates raises TFP by 1.2%. By contrast, Lange and Topel (2006) find more modest effects, and Acemoglu and Angrist (2000) find that a one year increase in schooling raises TFP by only 1 − 3%. In fact, Bedard (2001) finds that labor market evidence is consistent with a signaling model and inconsistent with a pure human capital model.

Given the uncertainty surrounding the magnitude of human capital externalities, I test two cases for the Lucas externalities: $\phi = 0.1$ and $\phi = 0.25$. In both instances, I recalibrate $\widetilde{A}$ to normalize steady state output to 1 when $\tau_a = 0$. Unsurprisingly, human capital externalities enhance the impact on wages of increased skill acquisition brought about by estate taxation. When $\phi = 0.1$, starting wages for high-signal workers increase by 3.07% in response to the tax compared to the more modest 1.89% increase in the baseline model. The effect is even more dramatic for the starting wages of low-signal workers, which rise by 4.31% instead of 0.58%. The reason in both cases is that higher endogenous TFP from the presence of additional skilled workers offsets the decline in the full information marginal product of labor caused by the lower capital stock. In fact, when $\phi = 0.25$, the increase in TFP more than completely reverses the decline in the marginal product of labor.

More intriguingly, the Lucas human capital externalities actually *mute* the increase in the skilled population from estate taxation. Instead of rising from 30% to 40%, as in the baseline, the skilled share rises by only 8% when $\phi = 0.1$ and by only 6% when $\phi = 0.25$. By buttressing the wages of both high-signal *and* low-signal workers, the human capital externality actually mutes the incentive to acquire skills. Overall, the combined effect of the information externality and the Lucas human capital externality is to increase output by 5% when $\phi = 0.1$ and by 7% when $\phi = 0.25$, compared to the 2% increase without the Lucas externality. To summarize, the magnitude of the human capital externality when $\phi = 0.1$ is comparable to the effect of removing the saving distortion, as done in the model without altruistic bequests, while the version with $\phi = 0.25$ gives substantially larger positive effects on output and welfare. Youth welfare responds more strongly to the Lucas externalities primarily because they receive higher wages regardless of skill level.

5.4 Optimal Estate Taxation

The results presented in the previous section demonstrate that estate taxation has a powerful effect on skill acquisition, wages, output, and welfare. However, that analysis assumed a 50% tax rate. In this section, I compute the *optimal* estate tax under different assumptions about signal accuracy. For $p = 70\%$, I also re-introduce the Lucas human capital externalities with $\phi = 0.1$. Furthermore, unlike in the previous section, I explicitly take into account the dynamic transition path between the initial and post-reform steady states when calculating welfare. In principle, policy changes that appear ex-ante optimal may no longer lead to welfare gains when transition costs are incorporated. The intractable debate over reforming Social Security in the United States partially reflects this dilemma.

To weigh both the benefits and costs of estate taxation, I numerically solve the social planner's problem for the optimal estate tax rate, assuming that revenues are used to finance targeted subsidies to new workers who receive a high signal. I assume that the social planner maximizes the average consumption-equivalent welfare change of all households in the economy. Specifically, the planner solves

$$\max_{\tau_a \in [0,1]} \int 100 \left(\left[\frac{W_x(a,\mu;\tau_a)}{W_x(a,\mu;0)} \right]^{1/1-\sigma} - 1 \right) d\Lambda_W(a,\mu) + \int 100 \left(\left[\frac{Y(a;\tau_a)}{Y(a;0)} \right]^{1/1-\sigma} - 1 \right) d\Lambda_Y(a),$$

where the first term represents the average consumption-equivalent welfare change for workers and the second term is the average change for youth.

One potential downside to this utilitarian objective function, however, is that it does not guarantee that a majority of households gain from a policy change. The possibility exists that large gains by a few households could lead to an average increase in welfare, even if a majority of households experience losses. To check whether this scenario occurs, I also report the fraction of households who experience strictly positive welfare gains in each case.

Table 6: Optimal Estate Rates

	$p = 60\%$	$p = 70\%$	$p = 80\%$	Lucas I*
Optimal Tax Rate	95%	30%	5%	70%
Δ *Welfare*	4.17%	0.46%	0.04%	2.18%
% *Who Gain*	95.3%	79.2%	65.9%	88.6%
Skilled Workers	53.0%	36.0%	31.1%	41.4%
Risk-Free Rate	4.31%	2.67%	2.10%	3.60%
ΔHigh-Signal Wage	6.11%	1.00%	0.29%	5.48%
ΔLow-Signal Wage	5.98%	−0.02%	−0.03%	4.11%
ΔOutput	7.52%	1.09%	0.40%	6.12%

*$p = 70\%$ in this case. For all versions, the initial steady is calibrated to match $\pi = 30\%$, $K/Y = 3$, and $Y = 1$.

Table 6 shows that the optimal estate tax is strictly positive and decreasing in signal accuracy. When $p = 60\%$, the planner optimally sets a *95%* tax rate, which helps 95.3% of households and leads to a utilitarian average welfare gain of 4.17%, *taking into account the*

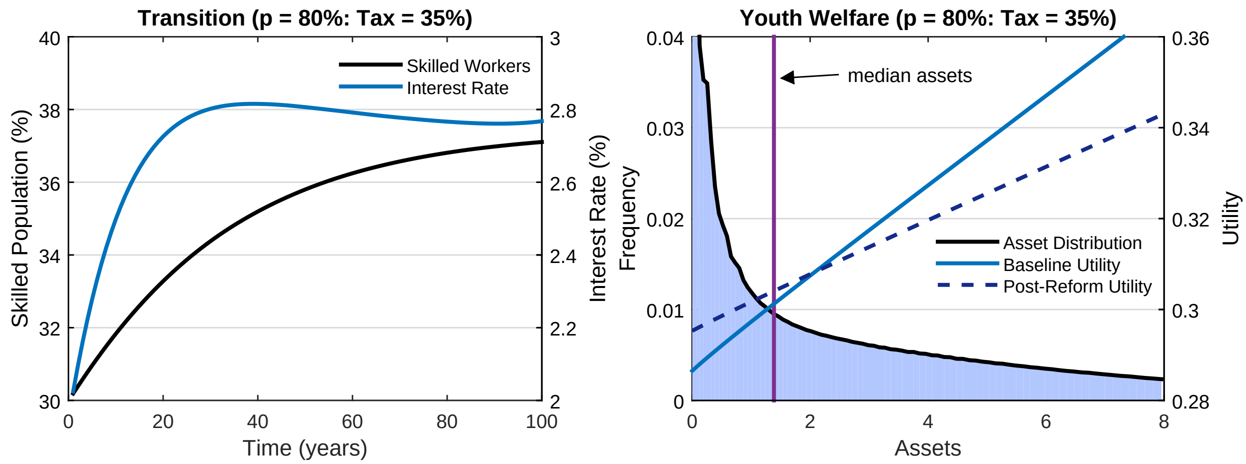

Figure 3: (Left) Transition dynamics of the skilled worker population and interest rates. (Right) Consumption-equivalent welfare before and after imposition of the estate tax. The shaded area is the initial pre-tax asset distribution.

transition path. By contrast, the optimal tax rate falls to 30% when $p = 70\%$ and to only 5% when $p = 80\%$. It is worth mentioning, however, that even when the labor market signal is highly accurate, estate taxation improves welfare for a large majority of households.

In fact, a majority of households experience welfare gains when $p = 80\%$ even at a much higher estate tax rate of 35%, as shown in figure 3. In this case, there is a utilitarian average welfare *loss* of -0.62%, but the loss is driven entirely by disproportionately large welfare losses among the wealthy. The panel on the right shows that the intersection of pre-reform and post-reform utility occurs to the right of median assets, which implies that a majority of households gain from the 35% estate tax.

Furthermore, these results likely underestimate the welfare gains because of the "perpetual youth" (i.e. no life cycle) and constant death probability assumptions in the model. When the estate tax is first implemented, there is an immediate jump in the share of youth who become skilled. However, when they enter the labor market, their higher skill proportion is diluted by the presence of many other older households who made their skill acquisition choices before the reform. If firms could observe workers' ages and the proportion of skilled workers *for each age*, they would condition wages by age. In this case, youth who acquire skills after the estate tax is implemented would experience large wage gains from the instant jump in the public signal π (rather than the gradual rise in π shown in the left panel).

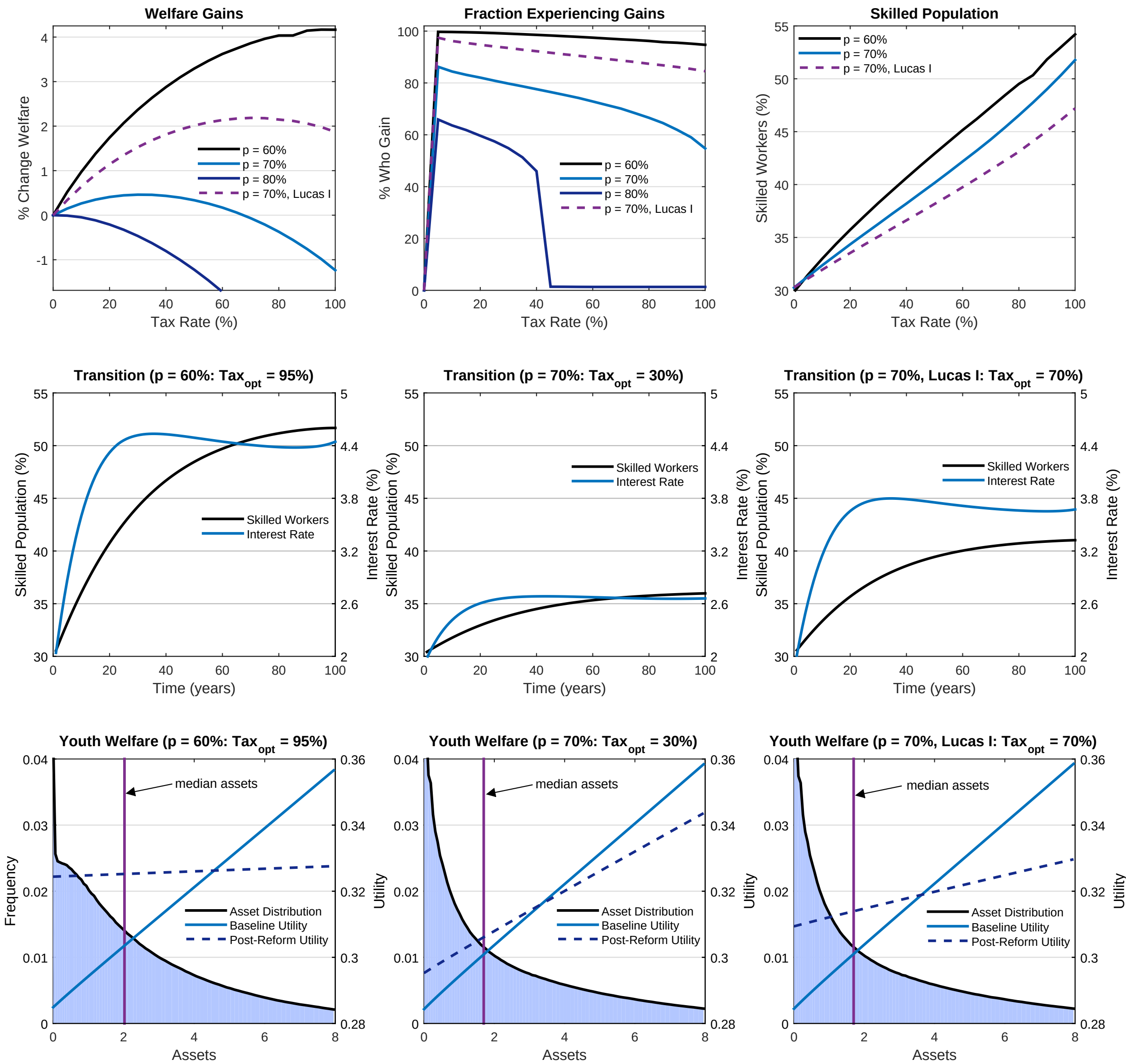

Figure 4: Welfare gains, transition paths, and welfare comparisons for the optimal estate tax for different signal accuracies, with and without Lucas human capital externalities.

Stepping back from the $p = 80\%$ case, figure 4 shows the results of the optimal estate tax exercise. In each instance, the optimal estate tax is an interior solution, $\tau_a \in (0, 1)$. Given the previous discussion of optimal estate taxation when $p = 80\%$, I focus on the remaining cases here. The first panel shows the average utilitarian welfare gain as a function of tax rate, while the second panel shows the percentage of households who experience gains. Notice that welfare improves for a majority of households *even with a 100% tax rate* (except when $p = 80\%$). However, large losses among the wealthy cause the planner to choose a lower rate to maximize its utilitarian objective function.

Consistent with findings in section 5.3, higher signal accuracy and the presence of Lucas human capital externalities both reduce the responsiveness of the skilled population share to changes in the estate tax, as shown in the top right panel. The middle row of panels plot the transition paths of the skilled population and interest rates in each case, and the bottom row of panels gives a visual representation of how the estate tax affects youth welfare. As expected, the policy is highly progressive, with poor youth experiencing the largest welfare gains and wealthy youth experiencing losses. However, a strong majority of households experience gains in all three cases.

6 Discussion

Section 3.3 ended by stating that theorem 3 could not establish whether an estate tax would benefit a majority of households or what the optimal rate would be. The quantitative results in section 5.4 answer the first question in the affirmative and demonstrate that the information externality created by the interaction of imperfect labor market signaling and endogenous skill acquisition is sufficiently harmful to overturn the familiar result of Chamley (1986) and Judd (1985). In particular, the dynamic model adds employer learning and endogenous capital accumulation to the static model—both of which attenuate the benefits of estate taxation—and still delivers optimal estate tax rates significantly above zero that deliver meaningful welfare gains for a large majority of households. Furthermore, dispensing with the assumption of perfect altruism or introducing Lucas (1988) human capital externalities both substantially increase the positive welfare gains of estate taxation.

Note that, contra the usual narrative about the Carnegie effect, parents in the dynamic model do not disapprove of or disagree with the effort decisions of their children. With perfect altruism, parents and their descendants display perfect time consistency and behave as one coherent infinitively-lived agent. Instead, the inefficiency in the model arises only from the effect that insufficient skill acquisition has on the public signal that employers use to help form beliefs over the skill level of each worker. Relaxing the assumption of perfect altruism simply reduces the sensitivity of parental capital accumulation decisions to the future taxation of the inheritances they pass on to their descendants. In such a world without altruistic bequests, estate taxation has even more power to improve welfare.

It turns out that shutting down the sensitivity of capital accumulation enhances the benefits of estate taxation about as much as does adding the Lucas human capital externalities. Even with learning, the skilled population share impacts long run wage dynamics through its use as a public signal. With Lucas human capital externalities, this skilled population share further enhances wages by contributing directly to higher endogenous TFP.

7 Conclusions

This paper investigates the effects of estate taxation when firms cannot directly observe worker skill levels. Imperfect labor market signaling gives rise to an information externality that causes agents to free ride off of others' skill acquisition. Specifically, agents do not internalize how their own skill acquisition impacts the skilled population share, which is a public signal that firms use in conjunction with individual signals to form beliefs over the skill level of each worker. Inherited wealth exacerbates the information externality because agents are risk averse and exert less effort the larger the inheritance they receive. By reducing these inheritances, an estate tax induces greater skill acquisition effort and increases the number of skilled workers. In the quantitative experiments, the optimal estate tax is significantly above zero, raises output and wages, and improves welfare for a large majority of households. Even with employer learning, the negative impact of an estate taxe on capital accumulation is more than offset by its salutary labor market effects.

These results provide avenues for additional research. Future work on optimal taxation and human capital policies should take into account how asset accumulation and information externalities influence human capital decisions. Furthermore, with richer labor supply behavior, estate taxes may play an even more positive role in tax reform by allowing the government to reduce distortionary labor taxes that also discourage skill acquisition. Apart from their implications for optimal taxation, information externalities have other policy implications worth exploring. For example, they represent an important consideration for the optimal design of safety net and welfare policies because of how unconditional transfers adversely affect skill acquisition that requires costly effort. Similar information externalities can also arise in other situations where agents benefit from and contribute to a group reputation. For example, grade inflation and financial aid may affect the degree to which college students shirk and free ride off the reputation of their alma mater.

References

Acemoglu, D., "A Microfoundation for Social Increasing Returns in Human Capital Accumulation," *Quarterly Journal of Economics* 111 (1996), 779–804.

Acemoglu, D. and J. Angrist, "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws," *NBER Macroeconomics Annual* 15 (2000), 9–59.

Bedard, K., "Human Capital versus Signaling Models: University Access and High School Dropouts," *Journal of Political Economy* 109 (2001), 749–775.

Bø, E. E., E. Halvorsen and T. O. Thoresen, "Heterogeneity of the Carnegie Effect," Working Paper, 2015.

Brown, J., C. C. Coile and S. J. Weisbenner, "The Effect of Inheritance Receipt on Retirement," *The Review of Economics and Statistics* 92 (2010), 425–434.

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

Elinder, M., O. Erixson and H. Ohlsson, "Carnegie Visits Nobel: Do Inheritances

Affect Labor and Capital Income?," *The B.E. Journal of Economic Analysis & Policy* 12 (2012), 1–37.

Fang, H. and P. Norman, "Government-Manded Discrimination Policies: Theory and Evidence," *International Economic Review* 47 (2006), 361–389.

Farhi, E. and I. Werning, "Progressive Estate Taxation," *The Quarterly Journal of Economics* 125 (2010), 635–673.

———, "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation," *Journal of Political Economy* 120 (2012), 398–445.

Gennaioli, N., R. La Porta, F. Lopez-de Silanes and A. Shleifer, "Human Capital and Regional Development," *Quarterly Journal of Economics* 87 (2013), 105–164.

Golosov, M., N. Kocherlakota and A. Tsyvinski, "Optimal Indirect and Capital Taxation," *Review of Economic Studies* 70 (2003), 569–587.

Holtz-Eakin, D., D. Joulfaian and H. Rosen, "The Carnegie Conjecture: Some Empirical Evidence," *Quarterly Journal of Economics* 108 (1993), 413–435.

Iranzo, S. and G. Peri, "Schooling Externalities, Technology, and Productivity: Theory and Evidence from U.S. States," *The Review of Economics and Statistics* 91 (2009), 420–431.

Joulfaian, D. and M. O. Wilhelm, "Inheritance and Labor Supply," *Journal of Human Resources* 29 (1994), 1205–1234.

Judd, K. L., "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics* 28 (1985), 59–83.

Kapicka, M. and J. Neira, "Optimal Taxation in a Life-Cycle Economy with Endogenous Human Capital Formation," Working Paper, 2013.

Krueger, D. and A. Ludwig, "On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium," *Journal of Monetary Economics* (Forthcoming 2016).

Lange, F. and R. Topel, "The Social Value of Education and Human Capital," *Handbook of the Economics of Education* 1 (2006), 459–509.

Lockwood, B., "Information Externalities in the Labour Market and the Duration of Unemployment," *Review of Economic Studies* 58 (1991), 733–753.

Lucas, R. E., "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22 (1988), 3–42.

Pestieau, P. and M. Sato, "Estate Taxation with Both Accidental and Planned Bequests," *Asia-Pacific Journal of Accounting and Economics* 15 (2008).

Piketty, T. and E. Saez, "A Theory of Optimal Capital Taxation," *Econometrica* 81 (2013), 1851–1886.

Popov, S. V. and D. Bernhardt, "Fraternities and Labor Market Outcomes," *American Economic Journal: Microeconomics* 4 (2012).

Rauch, J. E., "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities," *Journal of Urban Economics* 34 (1993), 380–400.

Restrepo, P., "Skill Mismatch and Structural Unemployment," Working Paper, 2015.

Shimer, R., "The Cyclicality of Hires, Separations, and Job-to-Job Transitions," *Federal Reserve Bank of St. Louis Review* 87 (2005), 493–507.

Straub, L. and I. Werning, "Positive Long Run Capital Taxation: Chamley-Judd Revisited," Working Paper, 2015.

A Supplementary Proofs

Proof of Lemma 1. The worker's objective function is continuous and defined over a compact domain, $[0,1]$. Therefore, the Extreme Value Theorem establishes that a solution exists.

Furthermore, the worker's objective function is strictly convex in e because of the strict convexity of v. As a result, $e(a;\pi,\tau)$ is single-valued in a and π.

Firms' beliefs $\mathbb{P}(x=1|s=h)$ and $\mathbb{P}(x=1|s=l)$ are continuous in π. Therefore, $w_h(\pi)$ and $w_l(\pi)$ are continuous in π. Because the choice set for e does not vary with a or π, and because the worker's objective function is continuous in e, a and π, an application of Berge's Maximum Theorem establishes that $e(a;\pi,\tau)$ is upper hemi-continuous. Because $e(a;\pi,\tau)$ is single-valued, it is also continuous.

■

Proof of Lemma 2. The only constraint that e must satisfy is linear, $0 \le e \le 1$, implying that the constraint qualification in the Kuhn-Tucker Theorem is satisfied, and thus,

$$(2p-1)[u(w_h(\pi)+a(1-\tau))-u(w_l(\pi)+a(1-\tau))]-v'(e(a;\pi,\tau))+\lambda_0-\lambda_1=0$$

where λ_0 is the Lagrange multiplier on the constraint $e \ge 0$ and λ_1 is the Lagrange multiplier on the constraint $1-e \ge 0$.

By assumption 4, $\lim_{e\to 1} v'(e)=+\infty$, implying that $e<1$ and hence $\lambda_1=0$.

The non-negativity constraint is also never binding, implying that $\lambda_0=0$. To see why, suppose that $\lambda_0>0$. Then $e=0$ and $\lambda_0=-(2p-1)[u(w_h(\pi)+a(1-\tau))-u(w_l(\pi)+a(1-\tau))]$ because $v'(0)=0$. Because $w_h(\pi)\le w_l(\pi)$ and $u(c)$ is an increasing function, $\lambda_0 \le 0$, which is a contradiction. Therefore, $\lambda_0=0$ and, furthermore, $e=0$ iff $w_h(\pi)=w_l(\pi)$.

■

Proof of Lemma 3. Recall that effort is determined implicitly by the condition $f(e,\pi,\tau;a)=$

0. Applying the Implicit Function Theorem and differentiating with respect to a gives

$$\frac{\partial e}{\partial a} = -\frac{f_a}{f_e} = \underbrace{\frac{(2p-1)(1-\tau)}{v''(e(a;\pi,\tau))}}_{>0}\underbrace{[u'(w_h + a(1-\tau)) - u'(w_l + a(1-\tau))]}_{<0} < 0$$

The strict convexity of v implies that the first term is strictly positive. The strict concavity of u and the inequality $w_h > w_l$ imply that the second term is strictly negative. Therefore, effort is decreasing in assets. Also, effort is clearly increasing in the estate tax rate because

$$\frac{\partial e}{\partial \tau} = -\frac{a}{1-\tau}\frac{\partial e}{\partial a} > 0.$$

■

Proof of Lemma 4. The condition $\phi_\pi(0,0) > 1$ ensures that a non-trivial equilibrium exists for $\tau = 0$ and therefore for all $\tau > 0$, because of the concavity of u.

The partial derivative of $e(a;\pi,\tau)$ with respect to τ at $(a;\pi(\tau),\tau)$ is given by the Implicit Function Theorem,

$$\frac{\partial e}{\partial \tau} = -\frac{f_\tau}{f_e} = \frac{(2p-1)a}{v''(e(a;\pi(\tau),\tau))}[u'(w_l(\pi) + a(1-\tau)) - u'(w_h(\pi) + a(1-\tau))]$$

The strict convexity of v implies that the denominator is strictly positive. Furthermore, the strict concavity of u and the inequality $w_l(\pi(\tau)) < w_h(\pi(\tau))$ imply that the numerator is also strictly positive. Thus,

$$\frac{\partial e(a;\pi(\tau),\tau)}{\partial \tau} > 0 \; \forall a \in [0,\overline{a}], \text{ implying that } \; \phi_\tau(\pi,\tau)|_{\pi=\pi(\tau)} > 0 \text{ and } \; g_\tau(\pi,\tau)|_{\pi=\pi(\tau)} > 0.$$

Now consider two tax rates, $\tau' > \tau$. Equilibrium $\pi(\tau)$ satisfies $g(\pi(\tau),\tau) = 0$. From the above result, $g_\tau(\pi(\tau),\tau) > 0$, there must exist some $\pi > \pi(\tau)$ such that $g(\pi,\tau) > g(\pi(\tau),\tau) = 0$.

From the Intermediate Value Theorem, $\exists \pi(\tau') > \pi > \pi(\tau)$ such that $g(\pi(\tau'),\tau') = 0$. In other words, higher tax rates cause a higher equilibrium proportion of skilled workers, $\pi(\tau') > \pi(\tau)$.

■

Proof of Theorem 4. By definition, $\mu^\tau \equiv \mathbb{P}(x = S|s^\tau) = \mathbb{P}(x = S|s^{\tau-1}, s)$. Applying Bayes' rule gives

$$\mu^\tau \equiv \mathbb{P}(x = S|s^{\tau-1}, s) = \frac{\mathbb{P}(s^{\tau-1}, s|x = S)\mathbb{P}(x = S)}{\mathbb{P}(s^{\tau-1}, s)}$$

Because s is independent from $s^{\tau-1}$, we can write this expression as

$$\mu^\tau = \frac{\mathbb{P}(s|x = S)\mathbb{P}(s^{\tau-1}|x = S)\mathbb{P}(x = S)}{\mathbb{P}(s^{\tau-1}, s)} = \frac{\mathbb{P}(s|x = S)\overbrace{\mathbb{P}(x = S|s^{\tau-1})}^{\mu^{\tau-1}}\mathbb{P}(s^{\tau-1})}{\mathbb{P}(s^{\tau-1}, s)}$$

Note that the ratio $\mathbb{P}(s^{\tau-1})/\mathbb{P}(s^{\tau-1}, s) = 1/\mathbb{P}(s|s^{\tau-1})$. Furthermore,

$$\begin{aligned} \mathbb{P}(s|s^{\tau-1}) &= \mathbb{P}(s \cap [x = S]|s^{\tau-1}) + \mathbb{P}(s \cap [x = U]|s^{\tau-1}) \\ &= \mathbb{P}(s|[x = S] \cap s^{\tau-1})\underbrace{\mathbb{P}(x = S|s^{\tau-1})}_{\mu^{\tau-1}} + \mathbb{P}(s|[x = U] \cap s^{\tau-1})\underbrace{\mathbb{P}(x = U|s^{\tau-1})}_{1-\mu^{\tau-1}} \end{aligned}$$

Because s is independent of $s^{\tau-1}$, this expression simplifies to

$$\mathbb{P}(s|s^{\tau-1}) = \mu^{\tau-1}\mathbb{P}(s|x = S) + (1 - \mu^{\tau-1})\mathbb{P}(s|x = U)$$

Substituting this expression above gives

$$\mu^\tau = \frac{\mu^{\tau-1}\mathbb{P}(s|x = S)}{\mu^{\tau-1}\mathbb{P}(s|x = S) + (1 - \mu^{\tau-1})\mathbb{P}(s|x = U)}$$

Lastly, recall $\mathbb{P}(s = h|x = S) = \mathbb{P}(s = l|x = U) = p$ and $\mathbb{P}(s = l|x = S) = \mathbb{P}(s = h|x = U) = 1 - p$. Thus, we can write

$$\mu^\tau = \frac{\mu^{\tau-1}p_s}{\mu^{\tau-1}p_s + (1 - \mu^{\tau-1})(1 - p_s)}$$

where $p_s \equiv p\mathbf{1}_{[s=h]} + (1 - p)\mathbf{1}_{[s=l]}$. ■