Ownership of the Means of Production

E. Glen Weyl* Anthony Lee Zhang†

March 2016

Abstract

Private property creates monopoly power. Common ownership can restore allocative efficiency, but it also destroys incentives for investments in the “capital” value common to all potential owners of an asset. Property rights should thus balance ex-ante capital investment and ex-post allocative efficiency through partial common ownership. A universal, self-assessed property tax with a universal right to force a sale at the self-assessed value implements this partial common ownership and is generically efficiency-enhancing compared to pure private ownership. At present, a range of calibrations suggests a 10% annual rate is robustly near optimal, implying expropriation of more than two thirds of capital income. However, the easiest application of our approach may be to assets with limited capital investment opportunities and administratively assigned property rights, such as radio spectrum or internet addresses, where taxes should be set at a higher rate equal to the (socially efficient) rate of annual asset turnover. This rate nonetheless retains some private ownership.

Keywords: property rights, market power, investment, asymmetric information bargaining

JEL classifications: B51, C78, D42, D61, D82, K11

---

†Stanford Graduate School of Business, 655 Knight Way, Stanford, CA 94305; anthonyzhang0@gmail.com.
What is common to the greatest number gets the least amount of care. Men pay most attention to what is their own; they care less for what is common; or at any rate they care for it only to the extent to which each is individually concerned.

– Aristotle, *The Politics*, Book XI, Chapter 3

Property is only another name for monopoly.

– William Stanley Jevons, preface to the second edition of *The Theory of Political Economy*

1 Introduction

Private ownership of the means of production is perhaps the oldest and deepest doctrine in mainstream economic thought, dating back to the Greek prehistory of the field and pervading contemporary thought. For example, Jacobs (1961) and de Soto (2003) argue the undermining or lack of property rights undermines investment incentives in rich and poor countries equally, leading Acemoglu and Robinson (2012) to consistently list property rights as the leading example of the “inclusive institutions” they argue foster economic development. On the other hand, economists such as Myerson and Satterthwaite (1981) have suggested that private property inhibits free competition, thus decreasing the efficiency of allocation, and conversely, Vickrey (1961) showed that under common ownership, full allocative efficiency can be achieved.

In this paper, we construct a simple model that incorporates both of these forces and implies a trade-off between private ownership that fosters investment and common ownership that fosters allocative efficiency. We show that a particular form of self-assessed capital taxation, proposed originally by Harberger (1965), continuously interpolates between private and common ownership, implementing optimally shared ownership. We calibrate our model and argue a 10% tax, extracting more than two thirds of capital income, is robustly near optimal for most capital assets that call for significant investments. However, in other cases, where investment is not important (e.g., radio spectrum) or where investment can be directly rewarded through objective assessment, the tax rate should be higher, equal to the socially efficient probability of per-period property turnover.

The role of property rights in incenting agents to invest in assets is an old idea in economics, recently explored in the theoretical literature by Grossman and Hart (1986) and Hart and Moore (1990). Although individuals have incentives for purely “selfish” investment (those that raise only that user’s value for using the good) under competitive common ownership (viz. Vickrey’s market socialism) (Milgrom, 1987; Rogerson, 1992), Che and Hausch (1999) show such schemes typically do not give individuals incentives to make investments that might benefit other potential users.\(^1\) In

---

\(^1\)Che and Hausch distinguish between “selfish” and “cooperative” investments. Our capital investment is the sum of a selfish and cooperative investment, because it augments the good’s value to both the investing and other users. Any of these types of investment may be synthesized with linear combinations of the other two, but in our context, the classification makes expositing the analysis simpler.
particular, common ownership offers zero incentive for an individual to make a “capital” investment that increases the value of the of the asset to all individuals equally, because such an investment raises the value of the good to the investing individual’s competitor as much as it raises the value to the investing individual. If the individual does not have an ownership stake in the good, such an investment is a pure waste from her perspective because any competitive mechanism with common ownership must be allocatively neutral in the face of such symmetric increases in all individuals’ values. Thus, some form of private ownership is crucial to encourage capital investment.

On the other hand, another strand of literature has argued private property is harmful to allocative efficiency. Myerson and Satterthwaite (1981) shows no mechanism can achieve fully efficient bargaining between a seller and buyer. Cramton, Gibbons and Klemperer (1987) shows this failure can be attributed to the structure of property rights – if both agents have partial ownership of the asset, efficient bargaining is possible.\(^2\) However, as Segal and Whinston (2013) discusses, no existing research combines these two perspectives, trading off the allocative efficiency costs from private property with the investment efficiency benefits. Furthermore, shared property rights are both complex to implement, and lead to efficiency only if fairly elaborate bargaining procedures are invoked (Segal and Whinston, 2011).

In this paper, we construct a model in which property rights affect the efficiency of allocation and investment, and show that a simple mechanism, employed historically albeit for very different purposes, can be used to trade-off these two distortions. A seller owns an asset; she first makes a costly investment to increase the capital value of the asset, and then makes a take-it-or-leave-it offer to sell the asset to a buyer. Partial common ownership of the property is implemented through a self-assessed tax: if the tax level is \(\tau\), the government owns a share \(\tau\) of all property.\(^3\) Hence, the seller must pay a fraction \(\tau\) of any sale price she proposes to the government, regardless of whether the buyer chooses to purchase the good. Because Harberger (1965) first proposed this form of self-assessed taxation (as a revenue-raising, rather than market-power-mitigating, device; see Subsection 5.1 below), we will refer to it henceforth as a \textit{Harberger tax}.

Raising the Harberger tax level increases allocative efficiency at the cost of investment efficiency. In the absence of a Harberger tax, allocation of the asset is inefficient due to the standard monopoly distortion – the seller will announce prices higher than her value for the asset. If the Harberger tax level is positive, the seller effectively becomes a buyer as well as a seller of the asset; thus, she will tend (Tideman, 1969) to announce lower prices, alleviating the monopoly distortion and increasing allocative efficiency. If the tax is equal to the probability of efficient sale – that is, the probability that the buyer’s value is higher than the seller’s – the seller will announce a price equal to her value for the asset, and the asset will be allocated fully efficiently. However, if the tax is nonzero, the seller pays taxes on the capital value of the asset. Hence, she does not appropriate the full marginal

\(^2\)This idea is further explored by Kaplow and Shavell (1996) and generalized by Segal and Whinston (2011).

\(^3\)As we discuss below, however, the appropriate interpretation of the share owned by the state is more ambiguous in a dynamic context.
benefit of her costly investment in the capital value of the asset, implying investment is suboptimal. The optimal Harberger tax level always lies between 0 and the probability of efficient sale, and its level depends on the elasticity of investment and the deadweight loss from market power.

Self-assessed taxes, besides being a tractable tool for modeling partial property rights, are straightforward to implement in practice. Every individual lists all property in a public cadaster and is taxed on the self-assessed value, but is required to sell any good listed there if offered that price. Similar systems have been used in a variety of settings dating as far back as ancient Rome, as documented by Epstein (1997). However, the closest argument we are aware of in previous work to ours that Harberger taxes could be used to improve allocative efficiency, rather than equitably collect taxes on valuations that are difficult to assess, was Tideman (1969)’s demonstration that Harberger taxes tend to increase the probability of sale. In a numerical example loosely calibrated to features of real-world housing markets, we show Harberger taxes near 10% significantly improve allocative welfare at little cost to investment welfare. Moreover, total welfare is fairly robust to small changes in the Harberger tax level about its optimal level – even Harberger taxes that are 5 percentage points larger or smaller than optimal achieve most of the welfare gains from the optimal Harberger tax.

Although 10% may appear a somewhat modest share for common ownership, we show that in a dynamic model (where we show such an annual rate is approximately optimal), it is actually quite radical, implying the expropriation of more than two thirds of capital income for reasonable discount factors. Furthermore, in some cases where our approach is most readily applicable, the tax rate should be substantially higher. When capital investment is a small concern, as with radio spectrum or internet names and addresses, the optimal tax converges toward the socially optimal, per-period rate of asset turnover. The same convergence also occurs when capital investments can be directly rewarded using objective assessments.

In Section 2, we construct our baseline static model, solve for the optimal tax level, and study the comparative statics of welfare with respect to tax level and of optimal taxation with respect to various primitives in a calibrated model. Because our proposal is most interesting in a dynamic context, where it is used to increase the frequency of asset turnover rather than simply to encourage a one-time sale, in Section 3, we study a dynamic extension of our model. We also use this model to show how to interpret our calibration in terms of measurable quantities per unit of time. As a result this analysis clarifies the radical magnitude of the economic change our optimal tax implies. In Section 4, we discuss various extensions, such as the effect of government observability of investment, the robustness of our results to alternative bargaining protocols, and relaxing our assumptions on the nature of buyers and sellers. In Section 5, we discuss our proposal’s relationship to previous capital taxation proposals, as well as various policies relating to intellectual property law. In Section 6 we discuss applications of our results to different forms of capital. We conclude in Section 7.

We present longer and less instructive calculations, proofs and calibration details in an appendix following the main text. Shortly we will post an online appendix with additional robustness checks.
2 Baseline model

In the baseline model, a seller holds an asset and makes a single take-it-or-leave-it offer to a buyer. In the absence of a Harberger tax, the seller is a monopolist, and announces a price higher than her value for the asset. If we impose a Harberger tax on the seller that is a fixed fraction of the price she announces, the seller effectively becomes a buyer as well as a seller of the asset, and thus has a lower incentive to overstate her valuation. When the Harberger tax rate is equal to the efficient probability that the good is sold, the seller’s optimal strategy is to announce her true value for the good. Thus, Harberger taxes can improve allocative efficiency; that is, the asset more often ends up with the individual with a higher value. However, imposing Harberger taxes on the seller decreases her incentives for making common-value investments in the asset, thus harming investment efficiency.

2.1 Setup

There is a seller $S$ who owns a single asset, and a buyer $B$. Utilities of $S$ and $B$ for the asset are, respectively,

$$v_S = \eta$$
$$v_B = \eta + \epsilon.$$

$\epsilon$ is a random variable representing heterogeneity in $B$’s value, which is not observed by $S$. $\epsilon$ is distributed according to $F(\cdot)$, which we assume is smooth and has full support on $\mathbb{R}$. Both agents are risk neutral; see Tideman (1969) for a partial analysis of the allocative problem that allows for risk aversion.

$\eta$ is a common-value component, which is a function of costly capital investment by $S$. $S$ chooses $\eta$, incurring a twice differentiable on $\mathbb{R}^{++}$ and strictly convex cost $c(\eta)$ to herself. We assume the seller has no idiosyncratic value simply for normalization; high idiosyncratic seller values are equivalent to a downward shift in the distribution of $\epsilon$.

For a given $\eta$, let $1_S, 1_B$ be indicators, which respectively represent whether $S$ and $B$ hold the asset at the end of the game, and let $t$ be any net transfer the buyer pays to the seller. Final payoffs for $S$ and $B$ respectively are

$$U_S = \eta 1_S - c(\eta) + t,$$
$$U_B = (\eta + \epsilon) 1_B - t.$$

Prior to the beginning of the game, the community decides on a Harberger tax level $\tau$. Then, $S$ and $B$ play a two-period game. In period 1, $S$ chooses $\eta$. In period 2, $S$ announces a price $p$ for the asset, pays taxes $p\tau$ to the cadaster, and then $B$ can decide whether to buy the asset by paying $p$ to $S$. The revenue the cadaster raises is distributed to the broader community in a manner we do not specify here, except to note the buyer and seller are a small part of this whole community, and
thus we neglect any impact the revenue raised has on their incentives.\(^4\)

### 2.2 Trade-off: allocative and investment efficiency

We solve the game by backwards induction. To begin, fixing \(\eta\) and \(\tau\), we will analyze behavior in the period 2 “taxed monopolist” game.

#### 2.2.1 Allocative efficiency

In the absence of a Harberger tax, \(S\) is a monopolist and will tend to announce a price higher than her valuation for the asset. However, if \(S\) pays a tax that increases with \(p\), she has an incentive to announce a lower price. \(S\)’s optimal choice of \(p\) balances these two incentives.

For any price \(p\), \(B\)’s optimal strategy is to buy the asset if her value is greater than \(p\), that is,

\[
\eta + \epsilon > p.
\]

Let \(m \equiv p - \eta\) be the seller’s *markup*. \(1 - F(m)\) is then the probability of sale when the markup is \(m\). The efficient probability of sale is \(q^{**} = 1 - F(0)\) We can invert this relationship to recover the inverse demand function \(M(q)\), which yields the markup that implies probability \(q\) of sale. Note \(M\) is strictly decreasing by our full support assumption. The seller’s expected utility is then

\[
\pi_S = M(q)q - \tau M(q) + (1 - \tau)\eta - c(\eta) = M(q)(q - \tau) + \kappa,
\]

where \(\kappa\) are terms which are independent of the choice of price/quantity \(q\) and depend only on the now-sunk investment \(\eta\). This *variable profit function* intuitively captures the trade-off between taxation and monopoly power. If \(q\) is larger than \(\tau\), the seller has an incentive to raise \(M(q)\) relative to the social optimum (by lowering \(q\)) so as to earn greater infra-marginal profits on the sales she makes. If \(\tau > q\), she has an incentive to lower \(M(q)\) to avoid taxation by raising \(q\).

Now suppose the Harberger tax is \(\tau = q^{**}\). If the seller chooses \(q = q^{**}\), she achieves 0 variable profits. However, for any other choice of \(q\), the seller will earn strictly negative variable profits. If she chooses \(q > q^{**}\), then \(M(q) < 0 < q - \tau = q - q^{**}\), whereas if she chooses \(q > q^{**}\), then \(q - q^{**} < 0 < M(q)\). Intuitively, if she sells with less than efficient probability, she becomes a net purchaser at an inflated price, but if she sells with more than efficient probability, she becomes a net seller at a deflated price. The best she can do, therefore, is to choose \(q = q^{**}\) and thus achieve efficiency.

Furthermore note that

1. Variable profits are strictly super-modular in \(q\) and \(\tau\) as \(M\) is strictly decreasing. Thus, for any \(\tau > q^{**}\), the seller will choose too large a probability of sale \(q^* > q^{**}\), and for any \(\tau < q^{**}\),

\(^4\)In fact, even this small issue can be addressed by simply raising the Harberger tax rate by ensuring no individual is paid marginally out of the pool of taxes she paid herself.
the seller will choose too small a probability of sale.

2. For any \( \tau \neq q^{**} \), the seller will choose a probability of sale strictly between \( \tau \) and \( q^{**} \). To see this, consider the case in which \( \tau < q^{**} \). If the seller chooses \( q = \tau \), then, because \( \tau < q^{**} \), \( M(q) > 0 \) and the seller has a strict incentive to increase \( q \) beyond that point. This observation implies a community that does not know the value of \( q^{**} \) can easily discover it: iteratively set \( \tau_{t+1} = q^*(\tau_t) \). Assuming each seller is small relative to the set of all sellers regulated this way, they will have a very small incentive to distort the government’s information acquisition process. Thus \( \tau \) will converge to \( q^{**} \).

These results are summarized in the following theorem.

**Theorem 1.** The probability of sale increases strictly in the tax rate. Allocative efficiency occurs if and only if the tax rate equals the efficient probability of sale. The probability of sale always lies strictly between the tax rate and its efficient level, so an iterative process of adjusting the tax rate to the current probability of sale will converge to the tax rate equaling the efficient probability of sale.

To consider the quantitative size of the distortion to allocative efficiency and the impact of \( \tau \) on it, note the seller’s first-order condition is

\[
M'(q)(q - \tau) + M(q) = 0,
\]

so that by the Implicit Function Theorem,

\[
\frac{\partial q^*}{\partial \tau} = \frac{M'(q^*)}{2M'(q^*) + M''(q^*)(q^* - \tau)} = \frac{1}{2 \left( \frac{M''(q^*)}{M'(q^*)} + \frac{M''(q^*)}{M'(q^*)} \right)} = \frac{1}{2 \left( \frac{M''(q^*)}{M'(q^*)} \right)^2},
\]

where the last equality invokes the first-order condition and drops arguments. Cournot (1838) showed this quantity equals the pass-through rate \( \rho(q^*) \) of a specific commodity tax into price; see Weyl and Fabinger (2013) for a detailed discussion and intuition. \( \rho \) is closely related to the curvature of the value distribution; it is large for convex demand and small for concave demand. It is strictly positive for any smooth value distribution and is finite as long as the seller is at a strict interior optimum. Myerson (1981)’s regularity condition is sufficient but not necessary for this second-order condition, as we show in Appendix A.1.

The marginal gain to social welfare from a unit increase in the probability of sale is equal to the gap between the buyer and the seller’s valuations, because the tax raised is simply a transfer. This gap is, by construction, \( M(q^*) \). Thus, the marginal allocative gain from raising \( \tau \) is \( M(q^*)\rho(q^*) \) or \( M\rho \) for short.

Note that \( M\rho \) is 0 at \( q^{**} \) so that no first-order social welfare gain results from additional taxation as we approach the allocatively optimal tax of \( q^{**} \). On the other hand, when \( \tau = 0 \), \( M\rho > 0 \) and a first-order welfare gain results from taxation.
2.2.2 Investment efficiency

Note the variable profits defined in the previous subsubsection were independent of $\eta$. On the other hand, the sunk profits, $(1 - \tau)\eta - c(\eta)$, depend on $\eta$. This sunk component, separably, determines investment incentives. In particular, a tax on self-assessed capital decreases the seller’s incentive to maintain the common value of her asset $\eta$. Because this component is entirely separable from the other component, regardless of what happens in the second stage of the game, the seller finds it optimal to equate:

$$c'(\eta) = 1 - \tau.$$  

We can define the following “investment supply” function $\Gamma(\cdot)$:

$$\Gamma(s) \equiv c^{-1}(s).$$

The value of the investment is always 1 on the margin, so the socially optimal investment is $\Gamma(1)$, whereas investment is only $\Gamma(1 - \tau)$ when the tax rate is $\tau$. By strict convexity of $c$, $\Gamma$ is strictly increasing, so the higher the tax, the more downward distorted the investment. We summarize this argument with the following theorem.

**Theorem 2.** Investment efficiency is achieved if and only if there is no tax, and the greater the tax, the more downward distorted is investment.

Again turning to the quantitative side, the increase in investment from a rise in $\tau$ is simply $\Gamma' = \frac{1}{\rho}$ by the inverse function theorem. The social value of investment is always 1 while the seller only invests up to the point where $c' = 1 - \tau$. Thus, the marginal distortion from under-investment is $\tau$. Thus, the marginal social welfare loss from raising $\tau$ is $\Gamma'\tau = \frac{\tau}{1 - \tau}\epsilon_\Gamma$, where $\epsilon_\Gamma$ is the elasticity of investment supply. Note that as $\tau \to 0$, this investment distortion goes to 0, so that there is no marginal investment distortion near the investment optimum of zero tax, whereas at any other tax (such as $q^{**}$), there is a strictly positive marginal investment distortion as long as $c$ is twice continuously differentiable in the relevant neighborhood and strictly convex.

2.3 Formula for optimal property rights

The socially optimal level of property rights balances the cost of taxation for investment efficiency with its benefit (below $q^{**}$) to allocative efficiency. It thus solves a classic optimal tax formula:

$$\frac{\tau^*}{1 - \tau^*} = \frac{M(q^*(\tau^*)) \rho(q^*(\tau^*))}{\Gamma(1 - \tau^*) \epsilon_\Gamma (1 - \tau^*)}.$$  \hspace{1cm} (1)

Under weak additional regularity conditions we discuss in Appendix A.1, this equation has a unique solution. The left-hand side is a monotone-increasing transformation of $\tau$ that appears frequently in elasticity formulas in the optimal tax literature; see, for example, Werning (2007). The
right-hand side is the ratio of two terms: the allocative benefit of higher taxes and the allocative distortion of higher taxes. The allocative benefit equals the product of the mark-up and the pass-through rate, whereas the investment distortion equals the product of the equilibrium investment size and its elasticity with respect to the keep-share $1 - \tau$. By the logic of the previous subsections, $\tau^* \in (0, q^{**})$.

**Theorem 3.** The optimal tax rate $\tau^*$ that maximizes social welfare is strictly positive and strictly below $q^{**}$ and is given by the (implicit, “sufficient statistics”) formula

$$\frac{\tau^*}{1 - \tau^*} = \frac{M \rho}{\Gamma \epsilon \Gamma}.$$

### 2.4 Calibration

In this section, we analyze a numerical example of Harberger taxation loosely calibrated to U.S. housing markets.

Our baseline model of Harberger taxation has three unknowns: the seller’s value for the asset; the distribution of buyer values, which determines the markup function $M(q)$; and the investment cost function $c(\eta)$. We will assume the distribution of buyer values is lognormal, with log mean normalized to 0, and the investment cost function is quadratic with cost $c(\eta) = \frac{\eta^2}{2g}$. These assumptions leave three parameters to determine: the value of the seller, which we will call $v_S$; the log standard deviation $\sigma$ of the lognormal; and the cost parameter $g$.

To match these parameters, we target three moments of real-world housing markets. Willekens et al. (2015) finds that, in the Danish housing market, the idiosyncratic component in house prices is lower bounded by 10% and upper bounded by 70%. We use the mid-point of this range at 40% of total housing value. Glaser and Shapiro (2003) suggest private ownership improves property prices by approximately 25% relative to rental. Also, using American Housing Survey data, Emrath (2013) estimate the average house is sold approximately once every 13 years. We will choose $\sigma$ so that the standard deviation of the lognormal distribution is 40% its mean value. We choose the cost parameter $g$ such that the total value of investment, which is $\frac{g^2}{2}$, is 25% of the expectation of the lognormal distribution. Finally, and given these calibrations, we choose the value of the seller $v_S$ so that the probability of sale under Harberger tax $\tau = 0$, the full monopoly regime, is $\frac{1}{13}$. Further details of the calibration can be found in Appendix C, and in the next section we show these annual interpretations in our static model are approximately correct in a dynamic model.

In Figure 1, we plot allocative, investment, and total welfare as a function of the tax rate. Increasing the tax from 0 sharply increases allocative welfare, at little cost to investment welfare. These figures represent our statements in Subsubsections 2.2.1 and 2.2.2 that investment welfare behaves quadratically near $\tau = 0$, whereas allocative welfare behaves linearly; hence, raising $\tau$ past 0 is always optimal. On the other hand, allocative welfare is quadratic about the allocative welfare-maximizing tax level, which is equal to the probability of efficient sale $q^{**}$. As a result, the slope of
allocative welfare with respect to $\tau$ is fairly low in a large neighborhood of $q^{**}$, whereas the slope of investment welfare is always larger for higher $\tau$, implying the optimal tax level is significantly below $q^{**}$.

The value of $\tau$ that maximizes total welfare is 0.11, and it improves total welfare by 0.8% relative to $\tau = 0$. Although the optimal $\tau$ is much lower than the efficient probability of sale, which in our calibration is $q^{**} \approx 0.2$, it achieves approximately 87% of total attainable allocative welfare gains, while incurring only 30% of the investment welfare loss of setting $\tau = q^{**}$. Moreover, total welfare about its optimum behaves basically quadratically, so total welfare is fairly robust to small changes in tax level near the optimal tax level – any $\tau$ within 0.05 of the optimal value achieves at least 80% of the total welfare gain of the optimal tax.

In Figure 2, we vary the three input moments about the values we use, and show how these variations change the total welfare gain, the percent of total allocative welfare gain achieved by the optimal tax, and the optimal tax level relative to the efficient probability of sale. Raising the standard deviation of the lognormal increases the importance of allocative welfare relative to investment welfare; hence, it increases the optimal tax relative to the probability of sale, as well as the percent of all possible allocative gains achieved at the optimum and the total welfare gain from the optimal tax. Raising the total value of investment does the opposite, decreasing the optimal tax, the total welfare gain, and the percent of all allocative gains achieved. Increasing the probability of efficient sale increases the optimal tax, but does so fairly slowly – for a seller whose value is lower than 70% of buyer values, the optimal tax is only approximately 25%. Because the allocative distortion is higher if the probability of sale is higher, increasing the probability of sale also increases the total welfare gain and the fraction of allocative gain achieved by the optimal tax.
Figure 2: Behavior of parameters with respect to input moments
Although these comparative statics move around most of the quantities significantly, interestingly, the optimal tax is quite stable across the reasonable calibration ranges we consider. It is never optimally much below 10% and never optimally much above 20%. Thus, 10% seems to be a quite robustly conservative and approximately optimal target. We show this conclusion is robust to many other factors in what follows.

3 Dynamic model

In our static model, trade can occur at most once and only one possible buyer exists. To extend our intuitions to a more general dynamic context and to investigate the robustness of our quantitative results to this context, we consider a simple dynamic model of repeated trade of a durable asset. To avoid dealing with repeated strategic interactions and thus potentially a large multiplicity of equilibria, our model is anonymous – each pair of agents interacts at most once, with the final owner of the asset remaining in the market and the other agent leaving forever. Moreover, our model assumes use values decay uniformly over time in a Markovian manner.

We show many of the intuitions from our static model extend to the general dynamic setting. In a case in which the government can set taxes contingent on the observed values of agents, setting the tax in each possible history equal to the efficient probability of sale in this history leads to fully efficient transfer in all periods. This connection between the dynamic and static model allows us to extend our static calibration to a simple dynamic context in which we can calculate the effect of a periodic Harberger tax on asset prices and the tax revenues collected each period.

3.1 Model

3.1.1 Preferences

Time is discrete $t = 0, 1, 2, \ldots \infty$. All agents discount utility at rate $\delta$. There is a single asset, which an agent $S_0$ owns at time $t = 0$. In each period, a buyer $B_t$ arrives to the market and bargains with the period-$t$ seller to purchase the asset through a procedure we detail in subsubsection 3.1.2 below. Hence, the set of agents is $\mathcal{A} = \{S_0, B_0, B_1, B_2, \ldots\}$. We will use $S_t$ to mean the period-$t$ seller who may, for example, be a buyer $B_{t'}$ from some period $t' < t$. In addition, we will often use $A_t$ to denote a generic agent in $\mathcal{A}$.

In period $t$, agent $A_t$ has utility $\gamma_{t}^{A_t}$ for the asset. The value of the first seller $\gamma_{t}^{S_0}$ and the values of all entering buyers $\gamma_{t}^{B_{t'}}$ are drawn i.i.d. from distribution $F$. Values evolve according to a Markov process: if any agent $A_t$ holds the asset at the end of period $t$, her use value in the next period $\gamma_{t+1}^{A_t}$ is drawn conditional on $\gamma_{t}^{A_t}$ according to the transition probability distribution $G(\gamma_{t+1} | \gamma_{t})$.

Assumption 1. $\gamma_{t} > \gamma'_{t}$ implies $G(\gamma_{t+1} | \gamma_{t})$ FOSD $G(\gamma_{t+1} | \gamma'_{t})$. 

11
This assumption states that agents with higher use values in the current period have uniformly higher use values in the next period, in the sense of first-order stochastic dominance. For our proofs, we will also assume $\gamma$ is bounded above almost surely by some constant $M$. However, we will relax this assumption in our numerical simulations. As above, we maintain quasi-linearity and risk neutrality.

### 3.1.2 Game

At the beginning of each period $t$, seller $S_t$ owns the asset. The use value $\gamma_{S_t}^t$ is drawn from $G(\gamma_{S_t}^S_t | \gamma_{S_t-1}^S_t)$ and is observed by $S_t$. Some tax level $\tau_t$ is set, which is known to $S_t$. Buyer $B_t$ arrives to the market, with use value $\gamma_{B_t}^t$, which is not observed by seller $S_t$.

As in the static game, $S_t$ can make a single take-it-or-leave-it price offer $p_t$ to buyer $B_t$. $S_t$ immediately pays taxes $\tau_t p_t$ to the government, and $B_t$ can decide whether to purchase the asset at price $p_t$. If $B_t$ purchases the asset, she pays $p_t$ to $S_t$, enjoys use value $\gamma_{B_t}^{B_t}$ for the asset, and becomes the seller in period $t+1$; that is, $S_{t+1} \equiv B_t$. Seller $S_t$ receives payment $p_t$ from $B_t$, and leaves the market forever. If $B_t$ does not buy the asset, $S_t$ retains ownership of the asset, enjoys use value $\gamma_{S_t}^{S_t}$, and becomes the seller in period $t+1$: $S_{t+1} \equiv S_t$. Buyer $B_t$ receives utility 0, and leaves the market forever.

### 3.2 Implementing allocative efficiency

We now consider how the socially optimal allocation can be implemented using a sufficiently rich Harberger tax. We begin by considering the social planner’s problem under complete information.

#### 3.2.1 Omniscient social planner’s problem

In each period $t$, the social planner chooses either $\{S_t, B_t\}$ to become the owner next period $S_{t+1}$. Her objective function is the discounted sum of use values of all agents. With slight abuse of notation, we can write the objective function as

$$V = \sum_{t=0}^{\infty} \delta \gamma_{S_t}^{S_{t+1}}.$$

A history $h_t$ describes all use values of all agents up to time $t$, including $\gamma_{S_t}^{S_t}, \gamma_{B_t}^{B_t}$. Because the problem is Markovian, the two values $\gamma_{S_t}^{S_t}, \gamma_{B_t}^{B_t}$ will in fact be sufficient for the social planner’s decision. A policy $\Phi (h_t)$ maps each history $h_t$ to an element $S_{t+1} \in \{S_t, B_t\}$; that is, for each history, the policy $\Phi (h_t)$ decides whether $S_t$ or $B_t$ will own and use the asset in period $t$, and thus become the seller $S_{t+1}$ in period $t+1$.

We will define the value function $V (\gamma_{S_t}^{S_t}, \gamma_{B_t}^{B_t})$ as the optimal value of $V$ from the start of $t$
onwards, if $S_t$ has $\gamma_{St}^t$ at the beginning of period $t$. The Bellman equation for the social planner is

$$V(\gamma_{St}^t, \gamma_{Bt}^t) = \max \left[ \gamma_{St}^t + \delta E \left( V(\gamma_{St}^{t+1}, \gamma_{Bt+1}^{t+1}) \ | \ \gamma_{St}^t \right) , \gamma_{Bt}^t + \delta E \left( V(\gamma_{Bt}^{t+1}, \gamma_{Bt+1}^{t+1}) \ | \ \gamma_{Bt}^t \right) \right].$$

**Theorem 4.** The solution to the social planner’s problem assigns the asset to whichever of $\{S_t, B_t\}$ has the higher use value $\gamma_t$ in every period $t$.

**Proof.** See Appendix B.

This policy is very simple: the good should always be allocated to the individual who obtains the highest use value at present.

### 3.2.2 Allocatively efficient tax policy

We now show the government can implement this policy through a sufficiently flexible Harberger tax, conditioned on the value of the seller $\gamma_{St}^t$ in each period.

**Theorem 5.** Suppose the tax $\tau_t(\gamma_{St}^t) = 1 - F_t(\gamma_{St}^t)$, for all $t$. Then, seller type $\gamma_{St}^t$ sells to all types with $\gamma_{Bt}^t > \gamma_{St}^t$ for all $t$, achieving the socially efficient outcome.

**Proof.** See Appendix B.

### 3.3 Investment

To incorporate investment incentives, we consider the simplest model we can calibrate in a plausible manner: one in which each period’s investment fully depreciates at the end of the period. In each period, prior to sale taking place, the current owner of the asset can invest $\eta$ at a cost of $c(\eta) = \frac{\eta^2}{2}$ to increase the value of the asset by $\zeta \eta$, for period $t$ only. $\zeta$ is the “global efficiency” of investment. The optimal level of investment is $\zeta$, generating value $\frac{\zeta^2}{2}$, and for any tax level $\tau$, per-period investment is $(1 - \tau) \zeta$. In Appendix A.2, we show that if investments are partially persistent, distortions increase for investments further in the future, in the sense that an investment that produces value only $t$ periods in the future is distorted by a factor $(1 - \tau)^{t+1}$. Because the qualitative features of the single-period investment model are similar to the general case, we focus on the single-period case for the remainder of our analysis.

### 3.4 Simplified calibration

To illustrate the basic features of our dynamic model with brevity, we consider a simplified extension of our static calibration in Section 2.4 rather than a calibration of the full model above.
Suppose that in each time period, entering buyers have use values drawn from a lognormal distribution, with log mean 0 and log standard deviation as in Section 2.4. At time \( t = 0 \), seller \( S_0 \) has use value \( \gamma^S \) for the asset, which is equal to the value of the seller in Section 2.4. As in the static calibration, total investment welfare is 25% of mean allocative welfare. We will specify a simple transition probability distribution: any owner of the asset in period \( t \), regardless of her use value, has value \( \gamma^S \) in period \( t + 1 \).

In this model, all sellers in all possible histories are identical. Hence, if the government sets some uniform tax level \( \tau \), sellers’ value function \( \gamma^S \) satisfies the Bellman equation:

\[
V_S = \max_{p_t} E \left[ \mathbb{1}_{WTP(\gamma^B_t) > p_t} (1 - \tau) p_t + \mathbb{1}_{WTP(\gamma^B_t) \leq p_t} (\gamma^S - p_t \tau + \delta V_S) \right]
\]

The value of any buyer for the good is her use value for the good in period \( t \), plus the constant seller value \( V_S \) in the next period: \( WTP(\gamma^B_t) = \gamma^B_t + \delta V_S \).

As in the static model, define the seller’s markup \( m = p - (v_S + \delta V_S) \). Because buyers and sellers have the same continuation value \( \delta V_S \), the continuation value does not affect the seller’s optimal markup. In Appendix A.3, we show the solution to the seller’s problem for any \( \tau \) is to choose a markup equal to what she would choose in a static problem in which her value is \( \gamma^S \) and buyer values are distributed as \( \gamma^B_t \). This result induces social welfare in each period equal to total social welfare in the corresponding static problem. Hence, the effect of taxes on welfare in the dynamic model are identical to those in our static model of Section 2.4, scaled up by a factor \( \frac{1}{1 - \delta} \).

In this dynamic model, we can also study the effect of changing \( \tau \) on the price of the asset and the tax revenues the government collects in each period. In Appendix A.3, we solve for the seller’s value function \( V_S \) as

\[
V_S = \frac{\pi(\tau) + (1 - \tau) \gamma^S}{1 - \delta (1 - \tau)},
\]

where \( \pi(\tau) = \max_q M(q)(q - \tau) \) is the optimal value of the monopolist’s profit under tax \( \tau \) in the static problem. Then, the asset price in each period is \( \gamma^S + m^*(\tau) + \delta V_S \), where \( m^*(\tau) \) denotes the optimal markup of a monopolist if the tax is \( \tau \), and the tax revenue collected each period is \( \tau \) times the asset price. If the markup is small relative to \( \gamma^S \), the asset price is approximately

\[
\frac{\gamma^S}{1 - \delta (1 - \tau)},
\]

which decreases approximately geometrically with increasing \( \tau \).

In Figure 3, we plot asset price and per-period tax revenue as a function of \( \tau \). The price of the asset declines quickly as we increase \( \tau \), and the total tax revenue collected as a fraction of the total use value of the asset in each period rapidly increases. For the optimal tax \( \tau^* = 0.11 \), total tax revenue collected in each period is 69% of the average use value of the asset in each period, causing the asset price to decrease by 68% relative to \( \tau = 0 \).
Intuitively, our dynamic model yields similar conclusions to the static model because, although it causes a much larger share of the total asset value to be expropriated by the same tax rate, it also increases dramatically the efficient *lifetime* turnover rate of the asset. These two effects cancel and yield essentially the same conclusion, though a somewhat different (and more radical) quantitative interpretation of the implied fraction of the good effectively commonly owned.

### 3.5 Discussion

Our model in Section 3.4 is very stylized. However, in Subappendix A.4, we numerically solve for equilibria of a dynamic model in which agents’ use values decay uniformly in percentile terms and show this alternative specification yields quantitatively similar results.

For our proofs, we impose the relatively strong assumption that values evolve in a Markovian manner. We have some evidence that this assumption is partly driving our results – in a more general model in which agents’ values may decay at different rates over time, we believe Harberger taxes may tend to be present-biased relative to the social optimum, preferentially transferring assets to individuals with higher-value decay rates. However, more complex models were much less tractable. We believe our model, although incomplete, captures most of the first-order features of trade in a dynamic environment, and we leave the study of Harberger taxation in a richer dynamic model to future research.

### 4 Extensions

In this section, we consider several extensions that investigate the robustness of our analysis in the baseline case as well as enrich it in various directions. For simplicity and because the two models
behave quite similarly from a qualitative perspective, all these extensions build off of the basic static model of Section 2, rather than the dynamic model of the previous section.

In Subsection 4.1, we show that if the government is able to observe and directly incent the seller to make common-valued investments, it can alleviate the investment efficiency losses from Harberger taxation, leading to higher optimal tax levels. Subsection 4.2 shows private-valued investments by the seller are efficient in our mechanism, regardless of the level of the Harberger tax. Subsection 4.3 shows that although Harberger taxes cannot achieve full allocative efficiency if sellers as well as buyers are heterogeneous, they can improve social welfare and achieve large fractions of total possible social surplus. In Subsection 4.4, we discuss bargaining protocols besides the simple one-sided offer mechanism we use. Finally, in Subsection 4.5, we show that increasing competition on the buyer side lowers the optimal level of the Harberger tax.

4.1 Partial observability

In our analysis above, the cadaster cannot directly observe investments made by individuals, and thus distorts investment when taxing self-assessed property. In some cases, though, the cadaster may be able to directly observe capital investments. After all, the leading property of capital investments is that they affect the objective value to all individuals and not just to the idiosyncratic value of the seller. Furthermore, for the purposes of imposing more traditional property taxes for real estate agents or other paid experts, making an objective appraisal of the market value of a property is common practice.

Beyond these common practices, a number of mechanisms could be used to elicit valuations from private agents, such as keeping cadastral values secret but having individuals offer bids on them that would be accepted upon exceeding cadastral value but that would also be used to estimate capital appreciation. Although such schemes could be gamed through collusion between potential buyers and the seller, law enforcement may be able to discourage such manipulations. Additionally, Levmore suggests other schemes for eliciting competing assessments, and the literature on mechanism design is continually developing more elaborate methods for elicitation in circumstances like these; see, for example, Crémér and McLean (1988).

If some mix of objective appraisal and such elicitation mechanisms could provide at least a noisy signal of capital value, the cadaster may be able to directly reward investments and thus avoid the investment distortion from the Harberger tax, which in turn raises its optimal rate closer to the allocatively efficient level. This argument captures some of Hayek (1945)’s intuition that local knowledge is what limits the prospects of common ownership. In this subsection, we formalize this intuition by following the analysis of Baker (1992) to construct optimal direct property subsidies to mitigate the moral hazard problem created by Harberger taxation.

As before, suppose $S$ chooses $\eta$ at cost $c(\eta)$. However, now suppose common value $\beta$ is deter-
mined by

\[ \beta = \zeta \eta, \]

where \( \zeta \) is a random variable representing the value of investment in different states of the world. There is local information about \( \zeta \); that is, \( S \) and \( B \) both observe \( \zeta \) prior to investment, but the cadaster does not.\(^5\)

The government can observe \( \eta \) and a signal \( \xi \). Prior to period 1, the government can choose an incentive scheme \( \gamma \psi (\xi) \), meaning that if the government observes \( \xi \), it will pay \( S \psi (\xi) \) for each unit \( \gamma \) of investment \( S \) makes. This policy is simply a negative property (property subsidy) based on an objective appraisal of \( \beta \).

Fix a realization of the signal value \( \xi \). If the government chooses reward function \( \psi (\xi) \eta \), and the Harberger tax rate is \( \tau \), investment level \( \Gamma (\psi (\xi) + (1 - \tau) \zeta) \) will be induced. For expositional simplicity, we now focus on the case when costs of investment are quadratic and thus \( \Gamma \) is linear: \( \eta (\zeta) = g\zeta \) for some \( g > 0 \), or, equivalently, that cost is \( c(\gamma) = \frac{\gamma^2}{2g} \). In Appendix A.5, we follow Baker and use a Taylor approximation to extend the analysis, approximately, to more general cost functions.

Given any choice of \( \psi (\xi) \), the seller chooses investment \( \Gamma (\psi (\xi) + (1 - \tau) \zeta) = g (\psi (\xi) + (1 - \tau) \zeta) \). Hence, for a fixed \( \tau \), the optimal \( \psi \) solves pointwise over realizations of \( \xi \) the maximization problem:

\[
\max_{\psi} E \left[ g (\psi (\xi) + (1 - \tau) \zeta) \zeta - \frac{(g (\psi (\xi) + (1 - \tau) \zeta))^2}{2g} \right] | \xi \].
\]

This program has the simple linear solution \( \psi (\xi) = \tau E (\zeta | \xi) \), which induces investment

\[
g (\tau E (\zeta | \xi) + (1 - \tau) \zeta) = g (E (\zeta | \xi) + (1 - \tau) (\zeta - E (\zeta | \xi))).
\]

Investment is thus equal to the conditional expectation of investment value conditional on the signal \( \xi \), plus a multiple \( 1 - \tau \) of the deviation \( \zeta - E (\zeta | \xi) \) from the conditional mean. For higher values of \( \tau \), investment is closer to the conditional mean.

The social welfare loss from this noisy estimation is

\[
IVL = \frac{\tau^2}{2} g E \left[ \zeta^2 - (E (\zeta | \xi))^2 \right] = \frac{\tau^2}{2} g (1 - r^2) Var (\zeta),
\]

where \( r^2 = E \left[ (E (\zeta | \xi))^2 \right] \) is the fraction of the variance in \( \zeta \) that is predictable by \( \xi \). If we take

\(^5\)An alternative formulation of this model is that the cost of investment is uncertain and is known to the seller and buyer, but the government only observes a noisy signal of the cost. Although this interpretation is more natural in many settings, we focus on the “investment value” interpretation to stay closer to Baker’s analysis and because it is simpler to present.
the derivative with respect to \( \tau \), we get
\[
\frac{dIVL}{d\tau} = -\tau g (1 - r^2) Var(\zeta).
\] (2)

As \( r^2 \) increases, \( \frac{dIVL}{d\tau} \) decreases in magnitude, and the socially optimal choice of the Harberger tax given by equation
\[
\frac{\tau^*}{1 - \tau^*} = \frac{M\rho}{\tau g (1 - r^2) Var(\zeta)}
\]
moves closer to the allocatively efficient level \( q^{**} \). Thus, to the degree that the cadaster can observe and reward capital investment, the detrimental effect of Harberger taxes on investment efficiency diminishes, and the optimal Harberger tax level is higher. This finding is consistent with Lange (1967)’s argument that the improvement of observation and computation through the improvement of information technology would increasingly make common ownership of the means of production feasible. In fact, ongoing work by a team of researchers led by Nikhil Naik, of which Naik et al. (2015) is a preliminary output, is attempting to automate high-quality property value assessment using Google’s Streetview images combined with computer vision and machine learning.

However, in contrast to Lange’s argument that such improvements in observability and computation would eventually make full common ownership (often described as socialism or communism) optimal, our analysis suggests that for any \( r^2 < 1 \), the optimal Harberger tax approaches only \( q^{**} \) and not 1. Thus even as information technology improves, fully common ownership may not be desirable.

However, note that our calculations in Subsection 3.4 above suggest a tax at the allocatively efficient turnover rate would likely expropriate the overwhelming majority of the value of private capital. For example, at a 5% discount rate, an allocatively efficient turnover rate of 30% would imply the expropriation of 85% of private capital value. Furthermore, for \( r^2 = 1 \), a full common ownership is efficient as well, and with heterogeneous sellers as we discuss in Section 4.3, it will typically be even more efficient than the optimal Harberger tax, so in a more general environment, and at the true limits of information technology or elicitation mechanisms described above, Lange’s argument may hold up in its simpler form.

### 4.2 Selfish investments

Thus far, we have assumed the seller’s investment only affects the common value of the good. Here, we show that if the seller can make pure private-value investments, that is, investments that affect only her value for the good, these investments are efficient, conditional on the final probability that the seller holds the asset. This efficiency guarantee is true regardless of the level of the tax, which is why we largely ignore such selfish investments in our analysis above.

Suppose seller \( S \) can pay cost \( c(\gamma) \) to increase only her own value of the good by \( \gamma \). Buyer \( B \)’s
value for the asset is $\epsilon$, unobserved by the seller and independent of $\epsilon$. Other features of the game are identical to those in Section 2.

As before, if $S$ announces price $p$, let $q = 1 - F(p)$ represent the probability that $B$ has a value higher than $p$. Fixing $\gamma$, $S$’s second stage payoff is

$$\pi_S (\gamma, \tau) = \max_p E [p 1_{c > p} + \gamma 1_{c \leq p} - p\tau].$$

In terms of the inverse demand function $M(q)$, profits equal to

$$\max_q qp(q) + \gamma (1 - q) - p\tau.$$

Let $q^*(\tau)$ represent $S$’s choice of $q$ for any given $\tau$. In the investment stage, $S$ chooses $\gamma$ to maximize $\pi_S (v) - c(\gamma)$. But, using the envelope theorem, we have that

$$\frac{d\pi_S (\gamma, \tau)}{d\gamma} = \frac{\partial}{\partial\gamma} [q^*(\tau)p(q^*(\tau)) + \gamma (1 - q^*(\tau)) - p(q^*(\tau))\tau] = 1 - q^*(\tau).$$

Hence, the first-order condition for $S$’s choice of investment is

$$c'(\gamma) = 1 - q^*(\tau).$$

This equation defines the constrained efficient level of investment, conditional on $S$ keeping the asset with probability $1 - q^*(\tau)$. Thus, Harberger taxation does not directly affect selfish investment efficiency. The only impact of Harberger taxes on such selfish investments is indirectly, through the probability of sale. For example, with absolute private property, selfish investments will tend to be too high relative to the social optimum, because $q^*$ is depressed below $q^{**}$. Thus, with pure private ownership, individuals will tend to invest in becoming excessively “attached” to their possessions relative to an optimal world where possessions turn over more frequently, as Blume, Rubinfeld and Shapiro (1984) first observed. However, this distortion occurs only because of the change in the turnover rate and not directly because of the tax. A few more informal observations are in order:

1. Our analysis shows that no issue of “hold-up” in the spirit of Grossman and Hart (1986); Hart and Moore (1990); Hart and Mooreq (1988) with purely selfish investments exists in this model. The reason is that the investor (the seller) makes a take-it-or-leave-it offer that the buyer thus cannot appropriate. This proeprty would not hold if the buyer made a take-it-or-leave-it offer. However, the very structure of the Harberger tax makes a seller offer the natural bargaining protocol, so we don’t consider this a serious concern.

2. Milgrom (1987) and Rogerson (1992) show that a Vickrey auction with common ownership or efficient bargaining through the similar Expected Externality mechanism (Arrow, 1979; d’Aspremont and Gérard-Varet, 1979) implies ex-ante selfish and privately observed invest-
ments are efficient. Our result indicates this result is driven entirely by the fact that these protocols lead to efficient allocations and not at all by the impact of these protocols on investment decisions conditional on investments, given that privacy ensures one’s bargaining partner cannot change his offer, even if the investor does not make a take-it-or-leave-it offer.

3. One objection to Harberger taxation is that individuals with high idiosyncratic utilities for properties must pay a high tax to avoid takings. To the extent that such idiosyncratic attachments are the result of random shocks an individual may experience, taxing away the benefits of such shocks could be either an advantage or disadvantage of Harberger taxation, depending on how this random gain in idiosyncratic utility interacts with the marginal value of consumption and thus whether individuals would like to insure against such shocks (as the Harberger tax effectively does). Regardless, our risk-neutral framework does not adequately capture these issues.

However, to the extent that such attachment is partially under the control of the owner, efficient Harberger taxation actually is an additional benefit of Harberger taxation. In particular, common experience suggests individuals tend to form much stronger attachments to objects that do not decay than to those that are durable, and to objects they own versus those they rent. Because of monopoly distortions, property tends to turn over much less frequently than is socially optimal. Thus, alleviating monopoly distortions may optimally lead individuals to become less attached to their possessions. This excessive attachment to material possessions is one of the principal flaws many religious and social thinkers perceive in capitalist societies; for a recent example, see Schor (1998). Therefore, unsurprisingly though heartening, moving optimally toward common ownership would tend to alleviate such attachment.

However, we note that, under a positive Harberger tax, $S$ has incentives to lower the value of $B$ for the asset. For example, with $\tau = 1$, lowering $B$’s value will allow $S$ to pay lower prices for the asset; thus, $S$ is willing to pay a cost to lower $B$’s value for the good, which is clearly socially inefficient. However, these kinds of “predatory investments” seem to us to be relatively rare compared to capital or selfish value investments; hence, we do not view the incentive for these as a major practical weakness of Harberger taxes. See Levmore (1982) for a more detailed discussion of predatory investment and how public monitoring could help prevent it.

4.3 Heterogeneous sellers

In our main analysis, we assumed seller values were known or followed a known dynamic process. If, on the other hand, seller values are heterogeneous, no single fully allocatively efficient Harberger tax exists. Different sellers will call for different optimal Harberger taxes; a seller with a high idiosyncratic value will efficiently sell infrequently, whereas a seller with a low idiosyncratic value will efficiently sell frequently. The community will have to set a Harberger tax ignorant of these
Let $\epsilon_S$ represent seller heterogeneity, normalized to have mean 0, and let $M$ be defined as in Section 2 for the case when $\epsilon_S = 0$. The first-order condition for allocative efficiency will have

$$\frac{dAW}{d\tau} = \int [M(\tau, \epsilon_S) - \epsilon_S] \rho(q(\tau, \epsilon_S)) dF(\epsilon_S) = 0.$$ 

The interpretation of this condition is that an allocatively efficient Harberger tax with heterogeneous sellers should be set so to balance the over- and underestimates of the efficient probability of sale for different values of the seller heterogeneity, with larger deviations (leading to values of $M - \epsilon_S$ further from zero) being penalized more, and with greater weight being given to states when the pass-through rate is high. It can easily be shown that any solution must be in the convex hull of the support of possible efficient sale probabilities. However, we did not find additional analytical results we could obtain very instructive, and thus focus the rest of this subsection on numerical analysis.

We repeat the calibration exercise of Section 2.4, allowing both seller and buyer values to be heterogeneous. We assume both seller and buyer values are distributed lognormally, with equal log standard deviations and possibly different means. The log standard deviation for both buyer and seller and the investment cost parameter $g$ are identical to those in Section 2.4. We normalize the mean log buyer value to 0, and then choose the mean log seller value such that the average probability of sale across all seller types under $\tau = 0$ is $\frac{1}{13}$, as before. We discuss further details of the calibration in Appendix C.2.
In Figure 4, we show allocative, investment, and total welfare as functions of $\tau$. The tax level that maximizes allocative efficiency is $\tau_{alloc} = 0.11$, which is lower than the average efficient probability of sale $q^{**} = 19\%$. The optimal Harberger tax level is 0.07; it improves total welfare by 0.3%, and it achieves 87% of the allocative welfare gain from the optimal tax. However, because buyers and sellers are heterogeneous, no Harberger tax can achieve full allocative efficiency – the optimal Harberger tax achieves only 37% of all possible allocative efficiency gains compared to a benchmark in which all welfare-improving trades happen. This imperfect targeting is why the optimal tax is lower and its induced welfare gains are lower. Thus although even in this case an optimal Harberger tax is useful and still leads to a rate close to 10%, substantially greater gains could be achieved by a cadaster that could condition tax rates on observables that correlate to the idiosyncratic value of a seller in the spirit of Akerlof (1978).

### 4.4 Other bargaining protocols and Vickrey subsidies

In the analysis above, we focused on one particular procedure for bargaining between sellers and buyers: sellers make take-it-or-leave-it offers to buyers. This procedure is simple, but in the environment above, it is clearly suboptimal. In our baseline environment with no uncertainty about seller values, if instead buyers were able to make take-it-or-leave-it offers to sellers, this would ensure full allocative efficiency. Thus, the concerns about allocative efficiency that motivate the Harberger tax for us might not really exist, invalidating our argument.

This reasoning, however, is an artifact of the assumption we made about the seller only having a known value. On the other hand, if sellers are heterogeneous, Myerson and Satterthwaite (1981) showed that with absolute property rights, no bargaining protocol can achieve efficiency. Furthermore, Segal and Whinston (2011), building on Cramton, Gibbons and Klemperer (1987), show that if property rights are shared between buyer and seller in proportion to the efficient probability of sale on average, full allocative efficiency always results for some feasible bargaining procedure. Although Harberger taxation does not achieve quite this optimum, in the previous subsection, we showed that Harberger taxes can achieve a large fraction of total achievable allocative efficiency even when sellers are heterogeneous. Furthermore, the mechanisms Segal and Whinston describe are fairly complex, because they involve allocating rights to potential buyers as well as sellers and involve fairly elaborate bargaining protocols. We view Harberger taxation as a simple mechanism that achieves most of the gains from more complex mechanisms. We leave the question of whether more sophisticated bargaining schemes could be made practical to future research.

Another arrangement that could avoid market power distortions and thus obviate the need for common ownership is an approximation to the “counter-speculation” subsidies Vickrey (1961) proposed. Although the exact Vickrey scheme requires common ownership as we discussed above, an approximation to the optimal scheme involves a simple subsidy on sales. In particular, a subsidy in the amount $\frac{M'(q^{**})q^{**}}{2}$ that is paid to the seller if and only if a sale takes place has the same effect as
as an allocatively efficient Harberger tax. Such a subsidy could conceivably be implemented without distorting investment incentives and thus could potentially be superior to an optimal Harberger tax.

We are concerned, however, that such a scheme would be impracticable for a variety of informational reasons. First, it requires knowing $q^{**}$ as with the allocatively efficient Harberger tax, but without a simple means to iteratively calculate the value, because it also depends on the value of $M'$. Second, $M'$ is particularly difficult to measure and requires a lot of cadastral authorities, especially given its value could be significantly context dependent in a way known to the seller but not to the cadastral authorities; see Weyl and Tirole (2012) for a detailed related discussion. Third, and perhaps most important, the scheme would be open to tremendous manipulation. Two-way sales could take place in succession and generate net subsidies to the participants. Finally, and perhaps most importantly, although this scheme would avoid common ownership in some sense, it would involve much more discretionary official intervention than would a Harberger tax. We therefore do not consider it a credible or less radical alternative.

4.5 Many buyers

Our analysis above assumes that there is a single potential buyer of the asset. In many realistic cases, several bidders may be competing to buy the asset. In this subsection, we analyze the effect of buyer-side competition on the optimal Harberger tax rate using a simple auction model.

Suppose the asset belongs to the seller $S$, and there are multiple buyers $B_1 \ldots B_n$. Suppose the values of the bidders are i.i.d. draws from distribution $F$. The asset is sold in a second-price auction, where $S$ can set a reserve price $p$. $S$ pays a tax on the reserve price $p$. Let $y_1$ represent the highest bid, and let $y_2$ represent the second-highest bid. $S$’s objective function is

$$
\pi_S = y_2 1_{y_1 > p} + p 1_{y_1 > p > y_2} + \eta 1_{y_1, y_2 < p} - p\tau.
$$

Taking expectations over $y_1, y_2$ and then taking derivatives with respect to $p$ yields

$$
\frac{d\mathbb{E} [\pi_S]}{dp} = P (y_1 > p > y_2) - \tau - m \frac{dP (y_1, y_2 < p)}{dp},
$$

where, as in Section 2, we define the markup $m \equiv p - \eta$. Substituting for the probability expressions, the derivative becomes

$$
\frac{d\mathbb{E} [\pi_S]}{dp} = nF^{n-1} (p) (1 - F (p)) - \tau - mnF^{n-1} (p) f (p).
$$

If we set this to 0, we get

$$
\tau = nF^{n-1} (p) (1 - F (p)) - mnF^{n-1} (p) f (p).
$$

(3)
Allocative efficiency is achieved when \( p = \eta \) and thus \( m = 0 \), which requires

\[
\tau = nF^{n-1}(\eta)(1 - F(\eta)).
\]

As \( n \to \infty \), this expression goes exponentially to 0. Thus, the allocatively optimal Harberger tax goes to 0 as competition grows. This conclusion is intuitive, given the follow-up to Jevons (1879)’s quote in our epigraph:

*But when different persons own property of exactly the same kind, they become subject to the important Law of Indifference...that in the same open market...there cannot be two prices for the same kind of article. Thus monopoly is limited by competition, and no owner, whether of labour, land, or capital, can, theoretically speaking, obtain a larger share of produce for it than what other owners of exactly the same kind of property are willing to accept.*

Larsen (2015) confirms this intuition empirically, and shows that sufficient competition in the market for used automobiles (if they are put up for an auction with dozens of bidders) limits the market-power distortion from property to 2%-4% of first-best allocative efficiency. Thus, in very competitive environments, or any environment where buyers have little idiosyncratic value for a particular piece of property, optimal Harberger taxes will be smaller than in an environment with greater market power.

Conversely, however, we have throughout that the buyer values the property under consideration on its own. If, on the other hand, that piece of property is complementary with many others, as is common in property development and the reassembly of spectrum (Kominers and Weyl, 2012b), then monopoly distortions substantially increase because no individual seller is pivotal in such a sale. This creates a “hold-out” problem (Mailath and Postelwaite, 1990). Introducing such complementarity would increase the optimal Harberger tax, and in the one case in which we are aware of Harberger taxation being proposed as a means of improving allocative efficiency, it was intended precisely to address such “eminent domain” issues (Tideman, 1969; Plassmann and Tideman, 2011). Thus, whether optimal Harberger taxes are above or below the levels we describe above depends largely on how large issues of complementarity or competition are relative to each other. We focus on a simple monopoly case as a compromise between these issues.

## 5 Connections

In this section, we relate our proposal to previous economic analysis and practices related to capital taxation and intellectual property.
5.1 Capital taxation

Although to our knowledge, the analysis of Harberger taxation as a means of lessening monopoly distortions is, with the exception of Tideman’s work on eminent domain discussed above, new to this paper, capital taxation is an ancient policy instrument. The central point distinguishing our argument from all previous proposals we are aware of is that it is directed only at improving allocative efficiency; it is not intended to raise revenue or redistribute endowments across the population.\(^6\)

The line of work closest to our proposal is the literature on market-based valuation schemes for taxation of assets whose value is difficult to ascertain. Various existing schemes for property taxation, such as the idea of taxing property based on previous sale value, are based on the idea of linking tax collection with observed market prices. For example, property taxes are often assessed based on previous transaction prices. This assessment method creates a variety of perverse incentives (e.g., to hold on to property when it rises in value and dispose of it when it falls) and does a poor job estimating the present property value. As a result, Harberger originally proposed self-assessment as a means of ensuring accurate assessments of tax liabilities for raising revenue. In fact, Levmore writes of the impact of self-assessment on turnover rates: “It is perhaps unfortunate that these side effects to self-assessment exist.” The author goes on to discuss methods of minimizing these “side effects.” This use of Harberger taxes has been criticized theoretically (Epstein, 1993) and empirically (Chang, 2012), primarily because it is unclear that self-assessed taxation schemes create incentives for truthful value revelation.\(^7\)

We essentially propose to invert the core argument of this line of work: rather than using information from market transactions to more effectively tax capital, we propose applying a tax on capital purely to increase the efficiency with which market transactions take place. Viewed in this light, the primary flaw of self-assessed taxation schemes discussed in Epstein (1993) and Chang (2012) – that proposed prices tend not to be equal to true values – is a core feature of our mechanism. At the welfare-maximizing level of the Harberger tax in our model, sellers are left with residual monopoly power, and announce prices higher than their values. However, this residual market power distortion is necessary in order to maintain optimal incentives for capital investment. In the context of our model, the purpose of self-assessed taxes is not necessarily to elicit truthful

\(^6\)As a result, the details of how the revenue raised is dissipated are relatively unimportant to our analysis. However, care must taken to avoid interfering with incentives: no individual should receive back a large fraction of the tax she pays herself.

\(^7\)Chang (2008) provides, to our knowledge, the only formal illustration of this argument, but the logic is clear from Theorem 1: the assessment is truthful if and only if the tax rate equals the efficient probability of sale, and any individual is entitled to take the property. In fact, in Harberger’s original proposal, any individual may take the property, but the tax rate was set to be equal to a usual property tax rate, on the order of a few percent, which is almost certainly far below the efficient probability of sale (it is below the probability of sale even in the present, which is distorted downward by market power). Thus, we would expect assessments at rates far above truthful values under Harberger’s system. On the other hand, Chang (2012) finds that in nearly all historical cases in which self-assessment has been used, only extraordinary actions (e.g., litigation or eminent domain takings) have triggered takings at the self-assessed value. In these cases, the probability of sale is far below the property tax rate, and thus we would expect, and Chang finds, that assessments are clearly below true values. To our knowledge, there is no example of Harberger’s system of universally invokable and universally applicable being implemented practically.
value assessments – for allocative efficiency to increase, it is enough that the value assessments are lower than the prevailing prices in a market absent the tax.

Various other rationales for capital taxation in the literature include capital as an inelastic base for redistributive taxation (George, 1879), local property taxes as a source of funds for local public goods provision (Lindahl, 1919; Bergstrom, 1979; Arnott and Stiglitz, 1979), a mechanism for dynamic redistribution (Judd, 1985; Chamley, 1986; Golosov and Tsyvinski, 2015), and a mechanism for governments without commitment power to avoid the temptation of appropriative redistribution (Farhi et al., 2012; Piketty, 2014; Scheuer and Wolitzky, Forthcoming). These arguments are largely orthogonal to our analysis, though in some cases, capital taxation using our mechanism can simultaneously accomplish some of the goals outlined in these works; in other cases, our tax will tend to create an excessive wedge along these dimensions that should be compensated by a subsidy (e.g., on savings) to ensure wedges are of optimal size. See, for example, our discussion of objectively assessed property subsidies in Subsection 4.1 above.

5.2 Intellectual property

Unlike physical property, intellectual property has always been limited in scope and duration precisely to limit monopoly power. In a sense, our argument is simply that intellectual and physical property should be treated more symmetrically.

However, unlike physical property, intellectual property is non-rivalrous in consumption, which implies the relevant activity distorted downward by property is not a turnover of the good from one rivalrous owner to another, but rather wide availability of the good.\(^8\) This difference in the nature of the distortion in turn means allocative efficiency can be achieved by simply setting the price of the property to zero. On the other hand, it makes rewarding investment much more complicated.

Given that our proposal is focused on reducing allocative distortions, it has less natural relevance to the problem of intellectual property. Conversely, we only considered the impact of Harberger taxation on a simple type of investment, namely, one in a uniform increase in values. By contrast, Weyl and Tirole (2012) argue that inventions differ along multiple dimensions in terms of the market for the products they produce. Although physical property usually has a tangible value and easily observable investments, the value of intellectual property is usually not apparent until a long process of marketing, adoption, and market testing has sorted out its value-added.

To make matters worse, charging a Harberger tax as a fraction of the total value of intellectual property requires knowing the total size of the market for the product if it were offered for free, so that the tax can be applied as a fraction of this total size. Unlike the probability of turnover, which is bounded between 0 and 1 and is plausibly in a knowable range for most goods, the value of market sizes will vary by orders of magnitude for observably similar products. For example, many apps on Apple’s App Store receive only a few downloads, whereas others “go viral” and are downloaded.

\(^8\)However, as Hopenhayn, Llobet and Mitchell (2006) point out, intellectual property rights may be rivalrous given that at most a single monopoly rent exists that must be divided among sequential innovators.
Without knowledge of this market size (which almost solves the problem itself, because a prize could be given directly), a Harberger tax would likely be laughably small for some markets while leaching all profits out of others.9

Kremer (1998) proposed a more promising and very different scheme that requires knowledge of the value created, on average, by the product. It may be realistic for the government to know this value than market size in some domains such as medicines. However, as Weyl and Tirole point out, even Kremer’s scheme can perform very poorly when, as is common in high technology, the average willingness to pay for a product and not just its market size are heterogeneous. They propose a more elaborate mechanism than ours, which would likely be more difficult to implement and require a more active discretionary state role. However, like us, they find that as public information becomes better, the optimal scheme converges to allocative efficiency and direct rewards for investment. In this sense, they arrive at a Langeian conclusion from a different starting point.

Finally, we can quantitatively compare the effects of Harberger taxes on asset value with the effects of existing limitations on intellectual property rights. Patent life in the United States is 17-20 years. Using discount factor $\delta = 0.95$, and assuming the value of any patent is constant for all time, the total value of intellectual property decreases by approximately 40% relative to a world with infinitely long patents. This figure is comparable to, though somewhat smaller than, our estimate in Section 3.4 that a 10% Harberger tax decreases asset values by approximately 70%. However, patent value likely decays over time in most cases, and patent scope and the probability a patent is enforceable (Lemley and Shapiro, 2005) are both limited by law, so these figures may effectively be reasonably close.

In any case, an important advantage of the Harberger tax over limitations on the length of ownership (such as applied to all property in Ancient Canaan according to the biblical account of the Jubilee) is that the monopoly distortion is convex in its size; thus, under weak conditions, a small reduction in it in each period is superior to limitation of its length. Gilbert and Shapiro (1990) formalized this idea in the context of intellectual property, but did not offer a concrete proposal for implementing such smooth limits on market power.

---

9If a public authority knew the efficient market size (call it $\sigma$) for a good, our scheme would be equivalent to setting a tax equal to $p\sigma\tau$, where $\tau$ is the Harberger tax as previously and $p$ is the price chosen by the monopolist. The required level of Harberger taxation to achieve taxation would be $\tau = 1$ (but would all eliminate monopoly profits and thus innovation incentive). A lower tax would still incent lower prices than pure intellectual property, but the innovator would be left with some rents, implementing the trade-off we analyzed above.

Even if such a scheme could be implemented, it would involve a substantially worse trade-off than with physical property, because the allocatively efficient tax is so high. However, it seems almost certainly impractical given the difficulty of estimating $\sigma$. For example, a $\sigma$ estimated at 1,000 would have no appreciable impact on the bottom line, and therefore prices of a product that ended up having a mass market. On the other hand, it would drive out all profits even at a modest 10% tax rate for a product with niche appeal to only a hundred clients.
6 Applications

In this section, we discuss two categories of applications of our approach. The first concerns goods for which capital investment plays a negligible role, and thus the primary focus is on allocative efficiency. For these goods, Harberger taxes seem particularly attractive, easy to implement, and perhaps even unlikely to raise substantial controversy. However, these goods are relatively limited, and thus we also discuss a broader implementation that could be applied economy-wide and would trade off investment incentives and allocative efficiency as in our calibrations above.

6.1 Purely allocative goods

Although many categories of property require substantial investments to maintain or improve, the value of some types of property is largely independent of investments. The leading example is radio spectrum, for which, at present, no known means exist to make permanent improvements to the spectrum itself, though individuals may make selfish investments in adapting themselves to broadcasting their programming on that band and may make some investment in marketing devices that tune well into that band. Other examples include internet names and addresses, some types of undeveloped land in rural areas, and certain natural resource extraction rights, especially for automatically self-renewing resources with little permanent temporal linkage. For the remainder of this subsection, we focus on the case of spectrum, but most of our arguments apply more broadly to these other cases.

In the early 1990s, the Federal Communications Commission (FCC) auctioned off most of the radio spectrum in an action that governments around the world have followed. The design of these auctions have been a subject of extensive study by economists, who played an important role in their design (Mirowski and Nik-Khah, 2007). Prior to this auction, standard practice had been to allocate spectrum by lottery and allow private bargaining in the spirit of Coase (1960) to reallocate spectrum to its most valuable use. As Milgrom (2004) emphasizes, a central motivation behind the design work was that monopoly distortions emphasized by Myerson and Satterthwaite (1981) stood in the way of this reallocation. The FCC and the economists it hired thought an auction could ensure an efficient allocation.

In practice, this proposal ran into two difficulties. First, designing a practical auction dealing with the full set of complexities in an environment like this (collusion, the combinatorial nature of bids, etc.) has proved daunting (Ausubel and Milgrom, 2005; Levin and Skrzypacz, 2014). Second, the auction procedure only deals with the monopoly distortions at one moment in time and does not facilitate the dynamic reallocation of property efficiently, thus necessitating significant design efforts to reassemble and reassign spectrum through another even more complicated auction process that is currently underway (Milgrom and Segal, 2015). We thus have reason to believe that if a more decentralized system free of the more severe monopoly distortions could be devised, it would significantly improve on the costly efforts now underway to repeatedly design auction mechanisms.
to address market power. Finally, dramatically reducing the value of spectrum licenses as capital, as a high Harberger tax would do, could significantly ease the capital constraints that Bulow, Levin and Milgrom (2009) show are an important barrier to efficiency in spectrum allocation.

One possibility would be to simply eliminate property rights and instead hold monthly auctions of some sort for rental rights on the spectrum. This approach would largely address the second problem, but the structure of each auction might still remain a challenge. Furthermore, this approach would obviously also undermine any investment incentives on spectrum. To the extent that concerns about such investments motivated the regime of property rights, a system retaining some investment incentives might be desirable, even if this concern is secondary to allocative efficiency. Thus, Harberger taxation, set at an allocatively efficient level, seems a natural policy option.

Empirical determination of allocatively efficient Harberger taxes is extremely simple, as we highlighted in Subsubsection 2.2.1: it simply requires iteratively setting the tax rate to equilibrium turnover rate. In spectrum markets, this approach might be a bit trickier because only a few nationwide spectrum owners exist, and thus they might have market power over the turnover rate that would give them strategic incentives to influence the tax rate. Historical and international data could likely be used to overcome this problem, though. Furthermore, given that spectrum property rights are administrative rather than absolute, this regime should be relatively straightforward to implement and would save on the substantial costs expended on auction design and participation. Spectrum thus seems the most natural place to begin experimenting with Harberger taxation.

6.2 Economy-wide implementation

In the previous subsection, we considered the simplest case for applying Harberger taxation. Now we consider how broad its scope should be in the long term. For the most liquid forms of currency and government bonds, Harberger taxation is neither harmful nor beneficial: these assets carry no market power, but also require no investment. One should therefore be indifferent to Harberger taxation of them except for the issues raised our discussion in Subsection 5.1 above about avoiding savings distortions. On the other hand, we positively advocate the application of Harberger taxation to essentially all other forms of capital, even very liquid ones, such as shares of companies. Although most individuals do not exercise market power over these commodities, some do, and these individuals also make decisions about investing their time in augmenting the value of these securities.

We see two broadly different ways to implement Harberger taxation in practice: one public and one private. The public system is most natural and accords with our discussion above. We only briefly describe how such a system would work here. A far more detailed exposition, including a variety of legal details, will appear in Posner and Weyl (Under Preparation). A new form of “sharing economy” platform would operate the private system, and we briefly describe how such a system could work.
Under the public system, all property would be registered in a cadaster with a regularly updated value. The cadaster would be made available, perhaps through a smartphone app, and the standard right of property would be replaced with a right to property that has not been purchased at the cadastral value, combined with a right to appropriate any property of another at its cadastral value. Cadastral proceeds would fund the enforcement of this system as taxes fund present-state enforcement of property law. Excess revenue could be returned to the community in any desired fashion, but a simple scheme would be to consider the cadaster as a common stock corporation whose profits would be returned as dividends to the public in proportion to share ownership. These rules provide a reasonably well-defined sense in which property would truly be held (partially) in common, rather than by some state actor meant to stand in for the common interest.

Note this system does not involve any greater centralized control than at present, except possibly for the maintenance of the technology for operating the cadaster. Individuals choose cadastral values and purchases. The cadaster enforces only the rules of the game, which are hardly more complicated than those involved in the absolute protection of property entitlements. The only discretionary element left to a planner is the level of the tax and the extent to which it differs across different categories of capital. However, as we showed above, a substantial but modest tax rate, such as 10% per annum, robustly improves over absolute property rights and achieves a large fraction of what a more tightly targeted tax could achieve. Given the incentives that a multiplicity of targeted tax rates would create for gaming, a relatively aggregated tax rate or even an entirely uniform one seems likely to be appropriate. However, for some categories of goods that are clearly distinguished, such as dwellings and vehicles, separate rates might be optimal. Additional complications, such as the treatment of property taken out of the physical jurisdiction of the cadaster, could be dealt with reasonably easily; for example, property removal could require the payment of a removal tax equal to the discounted present value of future tax flows.

Although we view this proposal as reasonably practicable, and note it could be implemented gradually by beginning with a very low tax rate that would rise only after its impacts are observed, any movement on such a broad issue of public policy is challenging. Therefore, considering whether implementing a system like this would be profitable for a private enterprise or platform, independent of broader social policy, would be worthwhile.

One possible such arrangement would be a broader or more radical version of various “sharing economy” platforms that have recently received significant commercial attention. A firm producing a particular durable commodity could, rather than offer it for traditional sale, sell goods only to members of the platform. Similarly, instead of selling goods entirely, the firm could sell only a right to use contingent on paying a Harberger tax, and offer the good for sale to any platform member.
at the self-assessed value. The platform would make revenue off the sales price, the Harberger tax, and the membership fee. As Armstrong (1999) shows, if this platform covered sufficiently many goods, setting the sale price and Harberger tax to maximize member surplus, which could then be extracted through the membership fee, would typically be in the interests of the platform. Although getting such a platform off the ground could present challenges, because it would have substantial network effects, dynamic pricing of membership might be able to overcome these challenges (Weyl, 2010).

To conclude, we briefly consider the magnitude of revenue in a world in which all capital is subject to Harberger taxation. Suppose, following our dynamic calibration of Section 3.4, the entire capital stock of the world is subject to a 10% annual Harberger tax. Following our calibrations, because such a tax would dramatically reduce the value of capital, it would generate annual revenue on the order of 3% or so of the aggregate capital stock as currently measured. According to Piketty and Zucman (2014), this capital stock is roughly five times gross domestic product in most advanced democracies. Combining these facts would imply revenue of approximately 15% of GDP, which is about two fifths of what the United States currently collects from all taxes. In this sense, the proposal is quite dramatic and “radical,” at least in terms of the change it would constitute in the flow of economic resources. Another way to see its radicalism is that it would, according to our calibrations, reduce the value of capital by about two thirds, bringing it back below its historically low levels in the mid-20th century, as documented by Piketty and Zucman.

As we discuss above, raising such large revenues need not lead to a large redistribution of those resources, because the revenue collected could be distributed in a manner specified by rules or given back as savings or property subsidies to offset other distortions. However, once such a large revenue stream is collected, a community would not necessarily wish to disperse it with the same inequality with which capital ownership itself is presently distributed. Because the Harberger tax offers a non-distortive (actually efficiency enhancing) means of generating a large pool of revenue, some communities might use it for redistributive purposes even if its justification is not redistributive.

The total welfare gains from our proposed policy are roughly 0.8% of the use value of the asset per period. Supposing these gains apply to roughly half the capital stock, which assuming a discount rate $\delta = 0.95$ generates flow value roughly 12% of GDP per year, would imply a total welfare gain of approximately 0.1% of national income per year, or about $20 billion per year.

7 Conclusion

In this paper, we argue the means of production should generally be owned neither in common nor privately, but rather through a mixed system that trades off the allocative benefits of common ownership against the investment incentives created by private ownership. We then show that a simple proposal for self-assessed capital taxes put forward by Harberger (1965) (for a very different purpose) can implement this system. Finally, we calibrate the optimal level of this tax in a variety
of settings, and we find a figure around roughly 10% per annum performs quite well and robustly in a range of settings.

Our analysis above considers only inanimate and not human capital. However, human capital receives a larger fraction of national income than inanimate capital and is likely as important a source of market power, given the unique talents many individual workers possess and the distortions to these talents, caused by labor income taxes. Indeed, most societies that have practiced common ownership of inanimate capital (e.g., the Israeli kibbutzim and the Soviet Union) have also socialized earning capacity to a significant extent. In these societies, human capital was largely directed according to social needs, rather than the choice of the human capitalist. Of course, these arrangements famously undermined human capital accumulation (Abramitzky and Lavy, 2014). Nonetheless, many methods exist for objectively assessing human capital that could be used to offer human capital subsidies to overcome this problem. In any case, partial common ownership would be a far smaller deterrent to investment than full common ownership. A fascinating question for future research is thus whether a workable system of more partial common ownership of human capital could be devised along the lines above. Such a system would have to deal with, among other challenges, the differing amenities of different workplaces (Sorkin, 2015) that make human capitalists far from indifferent across competing purchasers of their labor.

Our proposal does not, of course, exhaust the possibilities for mitigating market power by changing the nature of property entitlements. As we discussed in Subsection 4.4 above, Segal and Whinston (2011) show that with an appropriate bargaining procedure and ownership explicitly shared with potential future buyers rather than the public at large, greater efficiency gains than from Harberger taxation are possible. Additionally, our findings in Subsection 4.3 suggested a targeted Harberger tax, with rates that differ across different categories of sellers, might also dramatically outperform a uniform tax.

Whether such schemes are practical or could be simplified is an interesting question for future research. As Heller (2008) argues, the market power property creates is particularly socially costly in settings where complementary goods must be assembled. Although reasonable Harberger taxation would largely obviate this problem by forcing value revelation prior to the announcement of public projects and thereby prevent most holdout problems, Kominers and Weyl (2012) show other means of relaxing property rights may also significantly or completely eliminate these concerns while maintaining full investment incentives.

Our analysis assumed only the present owner of an asset may enjoy it and invest in it. In many circumstances, enjoyment and investment may be temporally shared in a variety of ways (Ostrom, 1990). Extending our analysis to such settings and devising appropriate forms of taxation in those contexts would be valuable. Another important practical issue absent from our analysis is transactions costs involved in transferring the possession of goods. Although we suggested in Subsection 6.2 that these costs may be diminishing with technology, they still pose significant challenges in many contexts, and any practical proposal would have to confront how they should
be born and how they should influence the frequency with which compulsory purchases should be allowed to occur.

Our analysis also assumed risk neutrality. This assumption may be reasonable when applied to relatively small capital goods or for wealthy individuals, but for many poorer individuals, this approximation is quite problematic regarding their attitude toward a house, which may represent a large fraction of their lifetime income. Tideman (1969) analyzes the Harberger-taxed-monopoly problem using a reduced form for risk aversion, but this approach does not allow him to study optimal taxation. An interesting extension of our analysis would allow for risk aversion.

Finally, although our calibrational analysis suggests a 10% tax is nearly optimal for many goods, and is a simple procedure to determine the allocatively efficient tax level for goods with small investment components (see Subsection 6.1 above), serious empirical analysis of the size of market-power distortions and investment elasticities for goods is crucial to pinning down this number with any precision. We hope future research will clarify these crucial elasticities as, for example, the literature on the elasticity of taxable income (Saez, Slemrod and Giertz, 2012) has clarified crucial elasticities for optimal redistributive taxation.
References


Appendix

A Analytical details and derivations

A.1 Baseline model

Here, we prove our statement in Section 2.2.1 that Myerson (1981)’s regularity condition is sufficient for \( \frac{\partial q^*}{\partial \tau} \) to be finite for all tax values below the efficient probability of sale \( q^{**} \). Myerson’s regularity condition states that marginal revenue is monotone. Revenue is \( M(q)q \). Taking a derivative yields

\[ M'(q)q + M''(q)q - \tau. \]

Taking the second derivative, we have

\[ 2M'(q) + M''(q)q - \tau < 0. \]

Now consider the monopolist’s problem under a Harberger tax \( \tau < q^{**} \). By Theorem 1, \( q(\tau) > \tau \); hence, \( 0 < q(\tau) - \tau < q(\tau) \). We want to show the following quantity exists:

\[ \frac{\partial q^*}{\partial \tau} = \frac{M'(q)}{2M'(q) + M''(q)(q - \tau)}. \]

So we have to show that the denominator is bounded away from 0. From our full-support assumptions on \( \epsilon \), \( M'(q) \) exists and is negative for all \( q \). If \( M''(q) \leq 0 \), we know \( q - \tau > 0 \), so \( M''(q)(q - \tau) \leq 0 \), and the numerator and denominator are both strictly negative; hence, their ratio is positive and nonzero and \( \frac{\partial q^*}{\partial \tau} \) exists. So suppose \( M''(q) > 0 \). Then

\[ 2M'(q) + M''(q)(q - \tau) < 2M'(q) + M''(q)q < 0 \]

Where we first use that \( 0 < q(\tau) - \tau < q(\tau) \), and then apply Myerson regularity. Hence, the denominator \( 2M'(q) + M''(q)(q - \tau) \) is strictly negative, and the ratio \( \frac{\partial q^*}{\partial \tau} \) exists and is positive.

Now we turn to regularity conditions for the social-maximization problem. From the text, the marginal benefit of increased Harberger taxation is \( M(q^*(\tau))\rho(q^*(\tau)) \) and the marginal cost is \( \Gamma'(1 - \tau)\tau \). Recalling that \( \rho = \frac{\partial q^*}{\partial \tau} \), the second-order condition for maximization is

\[ M'\rho^2 + \rho'M\rho + \Gamma''\tau - \Gamma'. \]

The first term is always negative (\( \rho > 0 > M' \)) and represents the “quadratic” nature of the allocative distortion discussed in the text. The final term is always negative as \( \Gamma' > 0 \) and represents the “quadratic” nature of the investment distortion. The two central terms are more ambiguous. However, Fabinger and Weyl (2016) argue \( \rho' \) is typically negative for most plausible demand forms (those with a bell-shaped distribution of willingness to pay, as we assume in most calibrations) and thus, given that \( M, \rho > 0 \), the second term is likely to be negative as well. The third terms is
genuinely more ambiguous. By the inverse function theorem, given that \( \Gamma = (c')^{-1} \),

\[
\Gamma'' = -\frac{c'''}{(c'')^3}.
\]

Assuming a convex cost function, this quantity is negative if and only if \( c''' > 0 \). Thus, a grossly sufficient condition (assuming \( \rho' \)) for the first-order conditions to uniquely determine the optimal tax is that \( c''' > 0 \). However, note this term is multiplied by \( \tau \), which is typically on the order of 10\% in our calibrations. Thus \( c''' \) would have to be quite negative indeed to cause a failure of the first-order condition to be sufficient.

### A.2 Dynamics and persistent investments

In our basic dynamic model and calibration, we have assumed investments in period \( t \) only affect value in period \( t \). Here, we consider a more general model in which investment can affect common values in future periods, and we show the effect of the investment distortion is exacerbated for periods further into the future.

Suppose that, in each period \( t \), \( S_t \) chooses an “investment parameter” \( \eta_t \in \mathbb{R} \) at convex cost \( c(\eta_t) \). For any period \( t \), the common use value \( \gamma_t \) for that period is determined by

\[
\gamma_t = \sum_{t'=\infty}^t H_{t-t'}(\eta_t),
\]

where \( H_t, t = 1, 2, \ldots \) are a collection of “investment effectiveness” functions that describe the value of a unit of investment \( \eta_t \) periods into the future. We will assume the following:

- \( \sum_{t=0}^\infty \delta^t H_t(\eta) \) exists for any \( \eta \): the discounted present value of investment is always finite.
- Each \( H_t(\cdot) \) is concave.

Now we consider \( S_t \)'s optimal choice of \( \eta_t \). First, because the problem is additive, we can ignore all \( \eta_{t'}, t' < t \) in determining the optimal choice of \( \eta_t \) in period \( t \). Moreover, in each period, the value function is linear in \( (1 - \tau_t) \gamma_t \). Thus, \( \frac{dV_t}{\gamma_t} = 1 - \tau_t \).

We can write the seller’s investment problem as

\[
\max_{\eta_t} w + (1 - \tau_t) H_0(\eta_t) + \delta (1 - \tau_t) V_{t+1} - c(\eta_t).
\]

Expanding the continuation utility,

\[
V_t = \max_{\eta_t} w_t + (1 - \tau_t) H_0(\eta_t) + \delta (1 - \tau_t) (w_{t+1} + (1 - \tau_{t+1}) H_1(\eta_t) + \delta (1 - \tau_{t+1}) \ldots) - c(\eta_t)
\]
\[ V_t = \max_{\eta_t} \left( \sum_{t'=t}^{\infty} \left( \prod_{t\leq t'} \delta (1 - \tau_t) \right) (w_{t'} + (1 - \tau_{t'+1}) H_{t'-t} (\eta_t)) \right) - c(\eta_t). \]

We can differentiate this expression with respect to \( \eta_t \), ignoring the \( w_t \) terms, giving solution:

\[ \sum_{t'=t}^{\infty} \left( \left( \prod_{t\leq t'} \delta (1 - \tau_t) \right) (1 - \tau_{t'+1}) H'_{t'-t} (\eta_t) \right) = c' (\eta_t). \]

In the case in which \( \tau_t = \tau \ \forall t \), this simplifies to

\[ \sum_{t=0}^{\infty} \delta^t (1 - \tau)^{t+1} H'_t (\eta_h) = c' (\eta_h), \]

whereas the social objective function is

\[ \max_{\eta} \sum_{t=0}^{\infty} \delta^t H_t (\eta) \]

with corresponding optimal solution:

\[ \sum_{t=0}^{\infty} \delta^t H'_t (\eta_h) = c' (\eta_h). \]

Hence, the distortion is higher further into the future, penalizing long-term common-value investments.

The intuition behind this result is as follows: Suppose the seller can make an investment that raises the common value of the asset \( T \) periods into the future. This investment will affect her value should she keep the asset until period \( T \), but she will also have to pay tax \( \tau \) on this value each period from the current period until \( T \). Hence, she only captures \((1 - \tau)^T\) of the value she creates by investment.

### A.3 Dynamic calibration algebra

Here, we solve the model of Section 3.4. The sellers’ value function \( V_S \) satisfies the Bellman equation:

\[ V_S = \max_{p_t} E \left[ 1_{WTP(\gamma_t^B_t) > p_t} (1 - \tau) p_t + 1_{WTP(\gamma_t^B_t) \leq p_t} (\gamma^S - p_t \tau + \delta V_S) \right]. \]

The value of any buyer for the asset is

\[ WTP (\gamma_t^B_t) = \gamma_t^B_t + \delta V_S. \]
Define the markup \( m = p - (\gamma^S + \delta V^S) \). We can write \( V^S \) as

\[
V^S = \max_{p^t} E \left[ \mathbb{1}_{\gamma^B_t > \gamma^S + \delta V^S} (1 - \tau^t) (\gamma^S + \delta V^S + m) + \mathbb{1}_{\gamma^B_t \leq \gamma^S + \delta V^S + m} (\gamma^S - (\gamma^S + \delta V^S + m) \tau + \delta V^S) \right]
\]

\[
= \max_{p^t} E \left[ \mathbb{1}_{\gamma^B_t > \gamma^S + m} (1 - \tau^t) m - \mathbb{1}_{\gamma^B_t \leq \gamma^S + m} \tau m \right] + (1 - \tau^t) (\gamma^S + \delta V^S).
\]

We can ignore the term \((1 - \tau^t) (\gamma^S + \delta V^S)\) in the optimization. Moreover, the term within the expectation depends only on \(\gamma^S\), and the distribution of entering buyer use values \(\gamma^B_t\), not on \(V^S\). Taking expectations of the indicators and changing variables to \(q = E \left[ \mathbb{1}_{\gamma^B_t > \gamma^S + m} \right]\), we can write

\[
V^S = \max_q M(q) (q - \tau) + (1 - \tau) (\gamma^S + \delta V^S).
\]

The maximization problem \(\max_q M(q) (q - \tau)\) is identical to that of the static problem in Section 2.4. Hence, in each period, sellers choose the same markup as they would choose in the corresponding static problem. Moreover, total social welfare is the discounted sum of welfare in each period; but welfare in each period is

\[
q E \left[ \mathbb{1}_{\gamma^B_t \geq \gamma^S + m} \right] + (1 - q) \gamma^S,
\]

which is, once again, equal to welfare in the corresponding static problem.

Let \(\pi(\tau) = \max_q M(q) (q - \tau)\) be the monopolist’s optimal variable profit, as a function of \(\tau\). Then, using that \(V^S = \pi(\tau) + (1 - \tau) (\gamma^S + \delta V^S)\), we can solve for \(V^S\):

\[
V^S = \frac{\pi(\tau) + (1 - \tau) \gamma^S}{1 - \delta (1 - \tau)}.
\]

Letting \(m^*(\tau)\) denote the optimal markup as a function of \(\tau\), the asset price is then the seller’s optimal price \(\gamma^S + m^*(\tau) + \delta V^S\), and the tax revenue collected in each period is \(\tau\) times the asset price.

### A.4 Dynamics with general transition probabilities

We consider an environment in which the government sets a single tax \(\tau\) for all time, and buyers and sellers whose use values evolve over time repeatedly trade an asset. The distribution \(F\) of entering buyer values is fixed, as is the transition probability distribution of values \(G(\gamma_{t+1} \mid \gamma_t)\). For any tax level \(\tau\), we seek a stationary equilibrium of the buyer-seller game. In any such dynamic game, the payoff of any seller type \(\gamma^S_{t^i}\) obeys the Bellman-like equation:

\[
V(\gamma^S_{t^i}) = \max_{p^t} E_{H_t} \left[ 1_{\text{WTP}(\gamma^S_{t^i} \mid p^t) > p^t} (1 - \tau^t) p^t + 1_{\text{WTP}(\gamma^S_{t^i} \mid p^t) \leq p^t} (\gamma^S_{t^i} - p^t \tau^t + \delta E \left( V(\gamma^S_{t+1} \mid \gamma^S_{t^i}) \right)) \right].
\]
The willingness to buyer of buyer $B_t$, type $\gamma_{t}^{B_t}$ is

$$WTP \left( \gamma_{t}^{B_t} \right) = \gamma_{t}^{B_t} + \delta E \left( V \left( \gamma_{t+1}^{B_t} \right) \mid \gamma_{t}^{B_t} \right).$$

Using this expression, we can evaluate the indicators in the seller’s value function:

$$V \left( \gamma_{t}^{S_t} \right) = \max_{p_t} E \left[ \left( 1 - \tau_t \right) \left( p_t \right) P_{\gamma_{t}^{B_t} \sim F} \left( \gamma_{t}^{B_t} + \delta E \left( V \left( \gamma_{t+1}^{B_t} \right) \mid \gamma_{t}^{B_t} \right) \geq p_t \right) \right] + \left( \gamma_{t}^{S_t} - p_t \tau_t + \delta E \left( V \left( \gamma_{t+1}^{S_t} \right) \mid \gamma_{t}^{S_t} \right) \right) P_{\gamma_{t}^{B_t} \sim F} \left( \gamma_{t}^{B_t} + \delta E \left( V \left( \gamma_{t+1}^{B_t} \right) \mid \gamma_{t}^{B_t} \right) < p_t \right).$$

The RHS of this Bellman-like equation depends on the value function $V$ in two ways: it affects the seller’s continuation payoff if she does not sell the asset, and it affects the distribution of buyer WTPs for the asset. As a result, this equation is not a true Bellman equation, and standard results such as the contraction property of the RHS may fail to hold.

We have been unable to prove the existence or uniqueness of a fixed point for this Bellman-like equation. However, in computational simulations, iteratively applying the RHS of the Bellman-like equation to a candidate $V$ function quickly converges to an apparently unique solution in all cases that we have tested. Hence, we have constructed a computational simulation of the single-tax-level, Markovian dynamic game, solving numerically for an equilibrium of the game for any given collection of parameter values.

We encountered numerical stability issues in attempting to solve the model matching the same moments as in Section 2.4, so we instead show results from a numerical example using a collection of parameter choices chosen to facilitate numerical computation, while maintaining nontrivial outcome comparisons across tax rates. The dynamic model has four main unknowns:

- Discount rate $\delta$
- Distribution of entering buyer values $F (\gamma)$
- Family of transition probability distributions $G (\gamma_{t+1} \mid \gamma_t)$
- Global efficiency of investment $\zeta$

We use discount rate $\delta = 0.9$. For the distribution of entering-buyer values $F$, we use a log-normal distribution with mean log 0, and some standard deviation $\sigma$. For the family of transition-probability distributions $G (\gamma_{t+1} \mid \gamma_t)$, we use a uniform-percentile-decrease model. An agent whose value $\gamma_t$ is at the $q$th percentile of all buyers in period $t$ has value $\gamma_{t+1}$ at the $\beta q$th percentile of buyers, where $\beta \in (0, 1)$. For the example shown here, we use $\sigma = 2$, $\beta = 0.9$, which leads to a 10% probability of sale with $\tau = 0$. Finally, we set the investment efficiency parameter to 20, which makes the magnitudes of allocative and investment welfare roughly comparable. We chose these parameters mostly to facilitate numerical computation, while maintaining nontrivial welfare comparisons across tax rates.
To numerically solve the model, we approximate the distribution of agent values $H_t$ using a 1,000-point uniform quantile grid. For each tax value on a 100-point grid, we iterate the RHS of the Bellman-like equation, which we will call $T[V]$, until convergence (defined as $\sup |T[V](\gamma_t) - V(\gamma_t)| < 10^{-6}$). Then, to evaluate the social welfare of a given equilibrium, we simulate the dynamic process by repeatedly calculating the optimal monopoly price for the current owner of the asset, drawing buyers from $H_t$, and transferring the asset to buyers if their values are higher than the monopoly price. We run each simulation trial for 5,000 time periods, and we run 2,500 simulation trials for each tax level. Reported allocative welfare levels are the average welfare of asset owners over all simulation runs.

In Figure 5, we show the asset price, tax revenue, allocative and investment welfare, and total welfare as a function of tax. Qualitatively, the results are similar to those of Sections 3.4 and 4.3. As in Section 3.4, the asset price declines approximately geometrically with the tax rate, and tax revenue quickly becomes a large fraction of total welfare. As in Section 4.3, the allocatively efficient tax is somewhat below the efficient probability of sale, and the optimal tax is lower than both. Although full allocative efficiency cannot be achieved with a single tax level, in this example, the optimal tax achieves over 80% of total achievable allocative efficiency gains.

### A.5 Partial observability

Here, we relax the assumption of Section 4.1 that costs are quadratic. As Baker (1992) shows, the following generalization of our results in Section 4.1 obtains. Assume $c(\eta)$ is a general convex function, such that $c''$ exists everywhere and is bounded away from 0. Furthermore, assume the conditional variance $E[(\zeta - E[\zeta | \xi])^2 | \xi]$ exists and is bounded above.

**Theorem 6.** If the variance of investment value unobserved by the government $E[(\zeta - E[\zeta | \xi])^2]$ is small, it is approximately optimal for the government to use the incentive scheme $\psi(\xi) = \tau E[\zeta | \xi]$. Under this scheme, the marginal gain in investment welfare $IW$ from changing $\tau$ is approximately

$$\frac{\partial IW}{\partial \tau} = -\tau E \left[ \frac{E[(\zeta - E[\zeta | \xi])^2 | \xi]}{c''(\hat{\eta}(\xi))} \right].$$

Because the government can choose arbitrary linear functions of $\eta$ as incentive schemes for each possible realization of $\xi$, we can solve the optimal incentive problem conditional on each possible realization of $\xi$. Suppose investment value is $\zeta \eta$, and the government observes signal value $\xi$. If the tax level is $\tau$ and the government chooses reward function $\psi \eta$, the investment level will be

$$\eta(\zeta, \psi, \tau) = c'^{-1}(\psi + (1 - \tau) \zeta) = \{\eta: c'(\eta) = \psi + (1 - \tau) \zeta\}.$$ 

Hence, fixing $\xi$, the government’s problem is to maximize investment welfare:
Figure 5: Generalized dynamics: tax revenue, asset price, and welfare vs tax
\[
\max_{\psi} E \left[ \xi \eta (\zeta, \psi, \tau) - c (\eta (\zeta, \psi, \tau)) \mid \xi \right].
\]

Take derivatives with respect to \(\psi\) to get, using Leibniz’ rule:

\[
\frac{\partial}{\partial \psi} E \left[ \xi \eta (\zeta, \psi, \tau) - c (\eta (\zeta, \psi, \tau)) \mid \xi \right] = E \left[ \frac{\partial}{\partial \psi} \left[ \xi \eta (\zeta, \psi, \tau) - c (\eta (\zeta, \psi, \tau)) \right] \mid \xi \right]
\]

\[
= E \left[ (\zeta - c' (\eta (\zeta, \psi, \tau))) \frac{d\eta (\zeta, \psi, \tau)}{d\psi} \mid \xi \right].
\]

Now, because \(\eta (\zeta, \psi, \tau) = c^{-1} (\psi + (1 - \tau) \zeta)\), we have \(c' (\eta (\zeta, \psi, \tau)) = \psi + (1 - \tau) \zeta\). Also, using the implicit function theorem, \(\frac{d\eta}{d\psi} = \frac{1}{c'' (\eta (\zeta, \psi, \tau))}\). Hence, the FOC rearranges to

\[
= E \left[ \frac{\psi}{c'' (\eta (\zeta, \psi, \tau))} \mid \xi \right] = E \left[ \frac{\tau \zeta}{c'' (\eta (\zeta, \psi, \tau))} \mid \xi \right]
\]

\[
\implies \psi = \tau \left( \frac{E \left[ \frac{\zeta}{c'' (\eta (\zeta, \psi, \tau))} \mid \xi \right]}{E \left[ \frac{1}{c'' (\eta (\zeta, \psi, \tau))} \mid \xi \right]} \right).
\]

In words, the FOC says \(\psi\) is equal to \(\tau\) times the weighted average of \(\zeta\), with weights \(\frac{1}{c'' (\eta (\zeta, \psi, \tau))}\). If the variance of \(\zeta - E [\xi \mid \xi]\) is small, \(\eta (\zeta, \psi, s)\) is nearly constant at some \(\tilde{\eta} (\xi, \psi, \tau) = \eta (E [\xi \mid \xi] \mid \psi, \tau)\) conditional on any \(\xi\); hence, the second derivative is also almost constant at \(c'' (\tilde{\eta} (\xi, \psi, \tau))\). Then, we can move the \(\frac{1}{c'' (\eta (\zeta, \psi, \tau))}\) outside the conditional expectation and cancel it, so the FOC rearranges to

\[
\psi (\xi) = \tau \frac{E \left[ \frac{\zeta}{c'' (\eta (\zeta, \psi, \tau))} \mid \xi \right]}{E \left[ \frac{1}{c'' (\eta (\zeta, \psi, \tau))} \mid \xi \right]} = \tau E [\xi \mid \xi].
\]

This proves the first part of Theorem 6.

Let \(\psi^* (\xi) = \tau E [\xi \mid \xi]\) represent the government’s optimal choice of \(\psi\) for each \(\xi\). To evaluate the social welfare achieved by the optimal choice of \(\psi\), we plug \(\psi^*\) into the government’s objective function:

\[
E \left[ E [\xi \eta (\zeta, \tau E [\xi \mid \xi], (1 - \tau)) - c (\eta (\zeta, \tau E [\xi \mid \xi], 1 - \tau))] \mid \xi \right].
\]

For notational simplicity, define \(\tilde{\eta} (\xi) = c^{-1} (E [\xi \mid \xi])\). Then, using the formula

\[
\eta (\zeta, \psi^* (\xi), \tau) = c^{-1} (\psi^* + (1 - \tau) \zeta)
\]

\[
= c^{-1} (\tau E [\xi \mid \xi] + (1 - \tau) E [\xi | \xi] - (1 - \tau) E [\xi | \xi] + (1 - \tau) \zeta)
\]

\[
= c^{-1} (E [\xi | \xi] + (1 - \tau) (\zeta - E [\xi | \xi])),
\]

47
we have from the implicit function theorem:

\[ \eta(\zeta, \tau E[\zeta | \xi], (1 - \tau)) \approx \bar{\eta}(\xi) + \frac{(1 - \tau)(\zeta - E[\zeta | \xi])}{c''}. \]

From a Taylor expansion of \( c \) about \( \bar{\eta}(E[\zeta | \xi]) \), we have

\[ c(\eta(\zeta, \tau E[\zeta | \xi], (1 - \tau))) \]

\[ \approx c(\bar{\eta}(\xi)) + c'(\bar{\eta}(\xi))(\eta(\zeta, \tau E[\zeta | \xi], (1 - \tau)) - \bar{\eta}(\xi)) + \frac{c''(\bar{\eta}(\xi))}{2}(\eta(\zeta, \tau E[\zeta | \xi], (1 - \tau)) - \bar{\eta}(\xi))^2. \]

Using again the Taylor approximation for \( \eta(\zeta, \tau E[\zeta | \xi], (1 - \tau)) - \bar{\eta}(\xi) \),

\[ \approx c(\bar{\eta}(\xi)) + c'(\bar{\eta}(\xi)) \left( \frac{(1 - \tau)(\zeta - E[\zeta | \xi])}{c''(\bar{\eta}(\xi))} \right) + \frac{c''(\bar{\eta}(\xi))}{2} \left( \frac{(1 - \tau)(\zeta - E[\zeta | \xi])}{c''(\bar{\eta}(\xi))} \right)^2 \]

\[ = c(\bar{\eta}(\xi)) + E[\zeta | \xi] \frac{(1 - \tau)(\zeta - E[\zeta | \xi])}{c''(\bar{\eta}(\xi))} + \frac{(1 - \tau)^2(\zeta - E[\zeta | \xi])^2}{2c''(\bar{\eta}(\xi))}. \]

Hence, plugging in both Taylor approximations to the government’s objective function,

\[ E[E[\zeta \eta(\zeta, \tau E[\zeta | \xi], (1 - \tau)) - c(\eta(\zeta, \tau E[\zeta | \xi], (1 - \tau))) | \xi]] \]

\[ \approx E \left[ E \left[ \zeta \left( \bar{\eta}(\xi) + \frac{(1 - \tau)(\zeta - E[\zeta | \xi])}{c''(\bar{\eta}(\xi))} \right) - \left( c(\bar{\eta}(\xi)) + E[\zeta | \xi] \frac{(1 - \tau)(\zeta - E[\zeta | \xi])}{c''(\bar{\eta}(\xi))} + \frac{(1 - \tau)^2(\zeta - E[\zeta | \xi])^2}{2c''(\bar{\eta}(\xi))} \right) | \xi \right] \right]. \]

We are interested in studying the behavior of this expression with respect to \( s \), so we can ignore the additive terms \( \zeta \bar{\eta}(\xi), c(\bar{\eta}(\xi)) \). Then, we are left with the term

\[ E \left[ \frac{1}{c''(\bar{\eta}(\xi))} \left( (1 - \tau) \zeta (\zeta - E[\zeta | \xi]) - (1 - \tau) E[\zeta | \xi] (\zeta - E[\zeta | \xi]) - \frac{(1 - \tau)^2(\zeta - E[\zeta | \xi])^2}{2} \right) | \xi \right] \]

\[ = E \left[ \frac{1}{c''(\bar{\eta}(\xi))} E \left[ (1 - \tau) (\zeta - E[\zeta | \xi])^2 + \frac{(1 - \tau)^2}{2} (\zeta - E[\zeta | \xi])^2 | \xi \right] \right] \]

\[ = \left( (1 - \tau) - \frac{(1 - \tau)^2}{2} \right) E \left[ \frac{E[(\zeta - E[\zeta | \xi])^2]}{c''(\bar{\eta}(\xi))} | \xi \right]. \]
Taking derivatives with respect to $\tau$, we have
\[
\frac{\partial}{\partial \tau} : -\tau E \left[ \frac{\left( \zeta - E [\zeta | \xi] \right)^2 | \xi \right] \frac{c''(\eta(\xi))}{c''(\eta(\xi))} \right].
\]

Hence, decreasing $\tau$ causes investment welfare to decrease, but at a rate proportional to the heterogeneity in investment value unobserved by the government $(\zeta - E [\zeta | \xi])^2$.

### B Proofs

**Proof.** [Proof of Theorem 4] The single-period value $\gamma_t$ is trivially increasing in value $\gamma_t$. This fact, together with our assumption 1 that higher $\gamma_t$'s imply uniformly higher $\gamma_{t+1}$, implies $V$ is component-wise increasing. For completeness, we sketch the fairly standard proof of this result. Similar arguments can be found in, for example, Stokey and Lucas (1989) and Smith and McCardle (2002).

Define the Bellman operator:
\[
T(W(\gamma^S_t, \gamma^B_t)) = \max \left[ \gamma^S_t + \delta E \left( W \left( \gamma^S_{t+1}, \gamma^B_{t+1} \middle| \gamma^S_t \right) , \gamma^B_t + \delta E \left( W \left( \gamma^B_{t+1}, \gamma^B_{t+1} \middle| \gamma^B_t \right) \right) \right].
\]

Because, by assumption, $\gamma_t$ is uniformly bounded above, this expression is a bounded discounted problem, and by standard arguments, $T$ is a contraction mapping with a unique fixed point.

Suppose $W$ is componentwise increasing in each component. Then, supposing $\gamma^S_{t+1} > \gamma_{t+1}$, by the FOSD property of $G$, we have $E \left( W \left( \gamma^S_{t+1}, \gamma^B_{t+1} \middle| \gamma^S_t \right) > E \left( W \left( \gamma^S_{t+1}, \gamma^B_{t+1} \middle| \gamma^S_t \right) \right)$. Hence, $\gamma^S_t + E \left( W \left( \gamma^S_{t+1}, \gamma^B_{t+1} \middle| \gamma^S_t \right) > \gamma^S_t + E \left( W \left( \gamma^S_{t+1}, \gamma^B_{t+1} \middle| \gamma^S_t \right) \right)$, and likewise for $B_t$ and $\gamma^B_t$. Hence $T(W(\gamma^S_t, \gamma^B_t))$ is componentwise increasing. Hence $V$, the unique fixed point of $T$, must be componentwise increasing.

Because $V$ is componentwise increasing and $G$ is FOSD-increasing in $\gamma_t$, we have that $\gamma^S_t > \gamma^B_t$ implies $\gamma^S_t + \delta E \left( V \left( \gamma^S_{t+1}, \gamma^B_{t+1} \middle| \gamma^S_t \right), \gamma^B_t \right) > \gamma^B_t + \delta E \left( V \left( \gamma^B_{t+1}, \gamma^B_{t+1} \middle| \gamma^B_t \right) \right)$. Hence, the social planner’s optimal strategy in each period is to assign the asset to the agent with higher $\gamma_t$. \hfill $\Box$

**Proof.** [Proof of Theorem 5]

We will first show that WTP for type $\gamma_t$ is increasing in $\gamma_t$.

Fix any distribution $H_t$ of WTP $\omega_t$ for buyers. The payoff of seller type $\gamma^S_t$ obeys the Bellman equation:
\[
V(\gamma^S_t) = \max_{p_t} E \left[ 1_{\omega_t > p_t} (1 - \tau_t) p_t + 1_{\omega_t \leq p_t} (\gamma^S_t - p_t \tau_t + \delta E \left( V \left( \gamma^S_{t+1} \middle| \gamma^S_t \right) \right)) \right].
\]
The interpretation of this equation is that, in each period $t$, seller $S_t$ is solving a “taxed monopolist” problem in which her reservation utility is $\gamma^{St}_t + \delta E (V (\gamma^{St}_{t+1}) | \gamma^{St}_t)$, and buyer values are distributed as $\gamma^{Bt}_t + \delta E (V (\gamma^{Bt}_{t+1}) | \gamma^{Bt}_t)$. Suppose the government commits to tax policy $\tau_t (\gamma^{St}_t) = 1 - F_t (\gamma^{St}_t)$. Then, the term $\gamma^{St}_t - p_t \tau_t (\gamma^{St}_t)$ is increasing in $\gamma^{St}_t$ for any $p_t$; that is, regardless of the price $S_t$ sets, her profits are higher if the tax she faces is lower. As in the proof of Theorem 4, together with the assumption that $G$ is FOSD-increasing in $\gamma^{St}_t$, this argument demonstrates $V (\gamma_t)$ is increasing in $\gamma_t$.

Because $V (\gamma_t)$ is increasing in $\gamma_t$, buyers’ WTP $\gamma^{Bt}_t + \delta E (V (\gamma^{Bt}_{t+1}) | \gamma^{Bt}_t)$ is also increasing in $\gamma_t$. To implement the socially efficient outcome – allocating the asset to the agent with higher $\gamma_t$ in every history – seller $S_t$ must set her price equal to her reservation utility $\gamma^{St}_t + \delta E (V (\gamma^{St}_{t+1}) | \gamma^{St}_t)$. As in the static case, this price is uniquely implemented by setting the tax equal to the efficient probability of sale; that is,

$$\tau_t = P \left( [\gamma^{Bt}_t + \delta E (V (\gamma^{Bt}_{t+1}) | \gamma^{Bt}_t)] \geq [\gamma^{St}_t + \delta E (V (\gamma^{St}_{t+1}) | \gamma^{St}_t)] \right)$$

$$= P (\gamma^{Bt}_t \geq \gamma^{St}_t) = 1 - F_t (\gamma^{St}_t).$$

Hence, under the tax path $\tau_t = 1 - F_t (\gamma^{St}_t)$, the asset is allocated in each history to the agent with higher use value $\gamma_t$; hence, the socially optimal allocation is achieved in every history. \qed

## C Calibration details

### C.1 Baseline model

As we discuss in Section 2.4, we aim to match three moments in the calibration:

- Standard deviation of lognormal divided by mean $PctSD$.
- Total value of investment divided by lognormal mean $InvVal$.
- Probability of sale under $\tau = 0$, $SaleProb$.

To match $PctSD$, fixing the log mean at 0, we choose the standard deviation $\sigma$ of the lognormal such that the standard deviation divided by the mean is equal to $PctSD$. Analytically, this condition is equivalent to

$$PctSD = e^{\sigma^2} - 1.$$

For the total value of investment, using $c(\eta) = \eta^2 2K$, the social welfare generated by investment at $\tau = 0$ is $\frac{K}{2}$. Hence, we set

$$K = 2 (InvVal) e^{\frac{\sigma^2}{2}}.$$
Finally, fixing a given lognormal distribution of buyer values, the probability of sale is a nonlinear decreasing function of the seller value. We numerically solve for a value of the seller that induces the probability of sale equal to $SaleProb$.

For any given tax level, we numerically solve for the seller’s optimal price, and then numerically integrate to find consumer surplus and thus total allocative welfare. We repeat this process on a 2,000-point grid of tax levels in $[0, 1]$ to produce the data for our figures.

In our online appendix (coming shortly), we repeat the calibration of subsection 2.4 using a range of different cost functions and distributions of buyer values, and show that the results are largely unchanged.

C.2 Heterogeneous sellers

To numerically solve the heterogeneous-sellers model, we use a grid approximation for sellers. For each tax level, we perform the single-seller numerical calculations described in section C for each seller type on a 200-point grid of evenly spaced quantiles of the seller value distribution, and then average quantities such as welfare and probability of sale across seller types.

In the online appendix, we replicate figure 2, showing the effects of varying input moments on welfare gains and the optimal Harberger tax level, in the heterogeneous-sellers setting. Additionally, we show that the calibration results are robust to using a range of different cost functions and distributions of buyer values.