Misallocations, Substitution,

and the Robustness of Activist Public Policy[*]

by Casey B. Mulligan

University of Chicago

January 2015

Abstract

This paper models productivity losses due to equilibrium misallocations of the factors of production, and public policies that might offset misallocations. Public policies in the model are implemented imperfectly in that a few underperformed activities are unintentionally taxed. A policy's impact on productivity is related to its degree of "caution," defined in much the same way that Friedman (1953) did in his study of countercyclical policies. Large baseline distortions and high substitution elasticities are associated with more caution. The paper also shows how the productivity impact can (or cannot) be measured with quantities versus prices.

Markets are sometimes inefficient. The actions of buyers and sellers may spill over to third parties (who are not connected to the transaction by contract). A seller may have market power that results in too little sales, or other market inefficiencies. The allocation of economic activity over the business cycle may feature booms and busts that are too volatile. Small businesses may suffer from imperfect access to credit markets. Neoclassical economics is replete with conceptual examples like these.

In theory, markets might be made more efficient by reducing barriers to trade. People and business that pollute might be forced to compensate the people harmed by their pollution. Antitrust policy might punish sellers for erecting market-entry barriers. Of course, actual public policy solutions have to deal with the realities of political economy, and the possibility that policies that reduce barriers may reduce efficiency by failing to go far enough (i.e., the theory of second best). These issues are treated well elsewhere in the literature, so this paper does not examine them in any detail. Instead, this paper models the impact of "indirect" public policies designed to serve much the same purpose as directly reducing barriers to trade.

Indirect approaches are routinely advanced in public policy analyses. For example, it may be proposed that small businesses be exempt from environmental regulations because they are already at a disadvantage in the marketplace. Or that the Federal Reserve "lean against the wind" over the business cycle, without necessarily directly addressing the fundamental problems that make the cycle too volatile. Labor unions may be given legal protections as an indirect means of promoting income equality. Another example of an indirect policy intervention: the Affordable Care Act's new subsidies for health insurance that, among other things, are intended to help correct the longstanding bias, originating in the income and payroll tax codes, against individual-market insurance in favor of employer-provided insurance.

My model of such policies assumes for the sake of argument a baseline economy with distorted markets. The baseline economy is a stylized version of Hsieh and Klenow's (2009) model, with distortions evidenced by low productivity: it is technically feasible to produce more

than the baseline economy with the same factors of production.[1] The subjects of this paper are public policy interventions that levy nonuniform taxes that are generally in the direction of taxing activities that are too prevalent in the baseline economy, and subsidizing the activities that are too scarce. But the correlation between an intervention's tax rates and the baseline distortions is not perfect.

Unless the correlation is far enough from zero, the tax policy reduces productivity even if it generally in the right direction. For this reason, my model might be described as a public finance version of Friedman (1953), who argued that imperfectly timed counter-cyclical macroeconomic policy might increase macroeconomic volatility, and Hansen and Sargent (2014). Manski (2005) is a public finance analysis, looking at optimal individual-level policy treatments with unknown and heterogeneous effects. My model also features policy-knowledge imperfections, but from a market perspective in which individual productivity depends on the treatments experienced by others in the marketplace. As discussed below, the paper also draws on behavioral public finance.

Section I sets up the model. Section II works out the relationships between policy interventions and productivity in the cases for which productivity has a simple variance representation. Section III considers the more general case. Section IV concludes with some comments on applications and rent-seeking activities.

---

[1] Restuccia and Rogerson (2008) also model misallocations among producers, but with a somewhat different analytical framework.

I.   A Model of Productivity and Misallocations

A. *Productivity*

The model features many ways to organize production on a real line $t \in [0,1]$. x_t denotes the input quantity of activity t. In order to discuss issues of productivity, I hold constant the total quantity of all of the inputs:

$$\int_0^1 x_t dt = 1 \tag{1}$$

The activities together produce value v for consumers according to a CES production function (2):

$$v = \left[\int_0^1 x_t^{\frac{\theta-1}{\theta}} dt \right]^{\theta/(\theta-1)} \tag{2}$$

where $\theta \in (0,\infty)$ is the constant elasticity of substitution between activities. At this level of abstraction, each activity may be interpreted as a separate good that enters a composite good v for the consumer, or as a particular use (e.g., location) of the limited factors of production in an overall production function of the form (2). Because total inputs are fixed at one, I refer to v as "productivity."

Because of the symmetry of the activities in technology, the productivity-maximizing allocation is $x_t = 1$ for all t and could, in principle, be decentralized with efficient trade in each production factor x_t at price $p_t = 1$.[2] But the baseline allocation of interest in this paper has market distortions, as represented by a baseline price or marginal rate of transformation profile p_t, with p_1 normalized to one, that is not the same for all t.

---

[2] In order to explicitly represent a situation in which the efficient allocation has unequally sized sectors, the interval $t \in [0,1]$ can be carved into unequal-sized pieces.

B. Distortions

Because it is so familiar from the literature, this paper skips an explicit derivation of how maximizing market participants substitute away from activities that are taxed and towards activities that are subsidized.[3] Equilibrium allocations can be summarized with a single equation (3):

$$x_t^{-1} = \int_0^1 \left(\frac{p_t}{p_s}\right)^\theta ds \tag{3}$$

As noted above, the baseline distortions can be interpreted as non-uniform taxes, with p_t greater for activities that are taxed and less for activities that are subsidized. But, as in Hsieh and Klenow (2009), the price or marginal rate of transformation profiles (hereafter, "price profiles") can represent varying degrees of market power, externalities, regulation, etc.

For any nonnegative elasticity of substitution, equation (3) shows that the allocation to any single activity t depends on the distortions of the other activities. As Adam Smith put it,

If … there was any employment evidently either more or less advantageous than the rest, so many people would crowd into it in the one case, and so many would desert it in the other, that its advantages would soon return to the level of other employments. (Smith 1776/1904, chapter I.10.1)

An activity-specific tax moves resources out of that activity until the extra marginal productivity there is enough to justify payment of the tax. Smith's result has been used repeatedly in modeling economic behavior in markets, but has received less emphasis in economic studies of policy robustness.[4] This paper argues that market equilibrium provides an additional reason why policy interventions may reduce productivity unless they are sufficiently "cautious."

---

[3] Hsieh and Klenow (2009) have a formal but accessible presentation.
[4] Public finance analyses of sectoral misallocations resulting from substitution effects include Harberger (1962), Shoven (1976), and Rosen (1986). Manski (2005) emphasizes issues of robustness, but primarily considers economic models in which any one agent's decisions is independent of the policy treatments experienced by other agents in the market. Hansen and Sargent's robustness models (Hansen and Sargent (2007) reviews several of them) are consistent with market substitution effects, but such effects are not emphasized in their presentation. Some of their models do feature strategic interactions among distinct market participants.

Of course, substitution behavior is not limited to "perfectly competitive" situations. There may be just one agent that has the right to engage in any one activity t, and thereby enjoys a degree of market power that varies inversely with the substitution parameter θ. The resulting rents are relevant for the rent-seeking discussion in Section IV, but are not relevant for the rest of the paper to the extent that the associated market power distortions are uniform across activities, as in Blanchard and Kiyotaki (1987), because any distortion that is common (in logs) to the activities cancels out in equation (3). A common distortion does affect aggregate factor supplies, but the purpose of this paper is to look at output for given aggregate factor supplies, which is why equation (1) holds them constant.

II.   Results for Small Baseline Distortions

The purpose of this paper is to compare baseline productivity with the equilibrium productivity associated with an alternative price profile p_t'. The log difference $\delta_t \equiv \ln(p_t'/p_t)$ can be interpreted as the incentive characteristics of a policy intervention imposed on the baseline economy. It is worth beginning with some results for baseline economies that are not too inefficient in that p_t is sufficiently close to one.

A. Variance-Correlation Representation

The productivity losses arise because distortions are not uniform across activities. Those losses are approximately summarized by the cross-activity variance of log price. The approximation is exact when the distortions are lognormally distributed:[5]

<u>Proposition 1</u> Lognormally distributed market distortions create log productivity losses that are proportional to the elasticity of substitution and cross-activity variance of the price profile.

---

[5] The lognormal distribution is most easily represented as $\ln p_t$ equal to in the inverse normal CDF evaluated at t, although another representation should be used when the sequencing of the activities is meaningful.

$$\ln v = -\frac{\theta}{2} var(\ln p) \equiv -\frac{\theta}{2} \left[\int_0^1 (\ln p_t)^2 dt - \left(\int_0^1 \ln p_t \, dt \right)^2 \right] \tag{4}$$

<u>Corollary</u>   In the neighborhood of efficiency, a policy intervention results in more productivity than the baseline economy if and only if it results in less cross-activity variance of the price profile.

The proof is in the appendix and shares the same arithmetic as portfolio theory (Merton 1969). However, variance here is not a measure of uncertainty, but rather a measure of heterogeneity. Production happens with full knowledge of all prices. But the productivity costs of deviations from efficient pricing are convex in the amount of the deviation.

For distributions that are not lognormal, equation (4) can be consider an approximation in the neighborhood of $p_t = 1$. However, the convexity of the deadweight costs of market distortions is not especially dependent on functional-form, or small-distortion approximations. Rather, the convexity of deadweight costs comes from the logic of market incentives: that executed market transactions are generating enough surplus to compensate participants for overcoming trade barriers. The first dollar of taxation in a sector eliminates transactions in that sector that would have been worth less than a dollar. The second dollar eliminates transactions that would have been worth somewhere between one and two dollars, and so on, with the marginal dollar of taxation doing more damage than any of the dollars that came before. Indeed, many of the results below require no restrictions on the magnitude of the substitution elasticity θ. In this regard, the logic of market distortions says a lot about the qualitative economics of policy interventions.

As modeled here, the deadweight costs of market distortions are second order, which might make them seem small compared with other sources of productivity changes over time and across countries. For example, if half of the economy were taxed forty percent more than the other half – something like the differential tax treatment of business capital and housing capital in the United States – the variance term in equation (4) would be about 0.065 and the

productivity loss would be about 0.033θ. Empirically estimating the productivity losses from misallocations is beyond the scope of this paper, but it is worth pointing out that prohibitions (or market distortions with many of the characteristics of prohibitions) of specific activities can result in large productivity losses in my model – approaching 100 percent as they would with $\theta \le$ 1. More important, as shown in this paper, the second-order nature of deadweight costs is the fundamental reason why it is "difficult" (defined rigorously below) for public policy to offset them rather than add to them.

The policy intervention that fully eliminates the baseline distortion, which I call the "precise policy," is $\delta_t = -\ln p_t$. Note that, for the precise policy, var(δ) = var(ln p). Any other policy intervention with var(δ) = var(ln p) I refer to as an intervention that is "tuned" to the baseline distortions, even though it does not exactly offset them (more on this below).

As discussed further in Section III, the precise policy is not always the one implemented. Policymakers may imperfectly understand the baseline distortions. Or perhaps the precise policy is well understood, but is politically infeasible. Under either interpretation, the practical policy choice may be between no policy (i.e., maintaining the baseline situation) or an imprecise policy. Define the correlation between the intervention and the baseline distortions as:

$$\rho(\ln p, \delta) \equiv \frac{\int_0^1 (\ln p_t) \delta_t dt - \left(\int_0^1 \ln p_t \, dt \right) \left(\int_0^1 \delta_t dt \right)}{\sqrt{var(\ln p) var(\delta)}} \tag{5}$$

Note that cross-activity correlation of log prices is the same as the cross-activity correlation of log quantities because, according to equation (3), log quantities are linear in log prices with slope $-\theta$ and intercept that is common to all activities. The correlation measures the tendency for increased distortion of activities that are underperformed in the baseline economy (a positive correlation) or decreased distortion of those activities (a negative correlation). The precise policy has $\rho = -1$.

B.   Friedman Revisited: A Negative Correlation is Insufficient

The impact $\Delta \ln v$ of a policy intervention δ on log productivity is, approximately:[6]

$$\Delta \ln v \approx -\theta \, var(\delta) \left[1 + 2\rho(\ln p, \delta) \sqrt{\frac{var(\ln p)}{var(\delta)}} \right] \qquad (6)$$

The productivity impact is negative unless the correlation is sufficiently negative. Any intervention (other than $\delta_t = 0$) with a zero correlation reduces productivity because, due to the convexity of deadweight costs, the productivity gains among activities with reduced distortions are more than offset by the productivity losses from activities among increased distortions. In other words, for the purposes of signing the productivity impact, it is not enough to know that the policy tends to increase distortions of activities that are overperformed in the baseline.

Equation (6) is reminiscent of Friedman's (1953) conclusion that countercyclical policies smooth economic time series only if they have effects on the time series that have a sufficiently negative correlation with the baseline business cycle. The critical correlation, he noted, was $-\frac{1}{2}$ in the case that the policy intervention was tuned to the baseline: $var(\delta) = var(\ln p)$ in my notation. That is my Proposition 2:

Proposition 2 If the baseline distortions are sufficiently small, or sufficiently close to normally distributed, then only the tuned policy interventions ($var(\delta) = var(\ln p)$) that increase productivity relative to the baseline are those that are precise enough that the correlation ρ between the policy distortions and the baseline distortions is less than $-\frac{1}{2}$.

Proof $\rho < -\frac{1}{2}$ is necessary for the right-hand side of equation (6) positive with $var(\delta) = var(\ln p)$.

Friedman obtained his 1953 results by positing a cost to variance of economic time series around a linear trend, which is an assumption he described as "somewhat arbitrary." Chetty (2001) notes that variance is a convex function of the deviations from trend, and that business cycle volatility might be measured in other ways: the absolute-deviation measure of volatility

---

[6] As noted above, the approximation is exact with a lognormal distribution of prices.

implies that any negative correlation would reduce volatility. Chetty concludes that the volatility measure relevant for welfare analysis of the business cycle is an "empirical question."

To the degree that a model like (1)-(3), which is consistent with any degree of substitution, is a good way of representing countercyclical policy and other policy interventions, convex costs is not an arbitrary assumption. Both Friedman's statistical model and the CES production function (2) weight activities equally in the sense that a given size and scope of the baseline distortion has the same productivity cost regardless of how it is allocated among the activities. Symmetry is a natural assumption for abstract discussion, where we have few specifics about the activities that are distorted by the various policies, but (in Friedman's words) is somewhat arbitrary. A nested CES is one way to allow for alternative substitution possibilities and, as shown in Section III, find critical correlations that differ from −½.

III.   Imprecise Policies and Results for Large Baseline Distortions

A. Policy imprecision

In order to say more about the reasons that a policy intervention might not coincide with the precise policy, and obtain results for large baseline distortions, take a stylized baseline economy with two categories of activities:

$$p_t = \begin{cases} e^\tau & if\ t \in [0,T] \\ 1 & otherwise \end{cases} \tag{7}$$

with $\tau > 0$. In words, the baseline allocation has T activities that are underperformed and the other $1-T$ activities overperformed. Equation (8) shows the public policy intervention to be compared with the status quo:

$$p_t' = \begin{cases} e^{\tau-\delta} & if\ t \in [0, T-m/2] \\ e^\tau & if\ t \in (T+m/2,1] \\ e^{-\delta} & if\ t \in (T-m/2,T] \\ 1 & if\ t \in (T,T+m/2] \end{cases} \tag{8}$$

9

$m \geq 0$ denotes the fraction of activities that are "misclassified" – that is, they are taxed when the precise policy would subsidize them, or vice versa. The misclassified activities are indicated in the bottom two rows of equation (8). If m were zero, then all of the activities would be properly classified and each activity would be taxed (subsidized) only if were overperformed (underperformed) in the baseline allocation. Figure 1 graphs the two price profiles (7) and (8).

[Figure 1 about here]

Recall that only relative prices matter and that p_1 is normalized to one. Therefore, Figure 1 and equation (8) do not assume that the policy intervention subsidizes some activities without taxing any others. In fact, the policy intervention could be achieved purely with taxes, by taxing activities $(T–m/2,T]$ and $(T+m/2,1]$ at the same rate and taxing the others at a zero rate.

Henceforth, a policy intervention refers to a pair (δ,m) describing the magnitude δ of the "taxes" and the prevalence m of misclassification according to equation (8). A tuned policy intervention has $\delta = \tau$. The precise policy intervention has $m = 0$ and $\delta = \tau$, and results in maximum productivity.

Under the policymaker-error interpretation of the model, $m > 0$ represents classification errors on the part of the policymaker. Both directions of misclassification are widely acknowledged in antitrust economics and other applications at the intersection of law and economics.[7] As the Supreme Court put it, "It is sometimes difficult to distinguish [true] competition from conduct with long-term anticompetitive effects," and that some enforcement actions will "chill the very conduct the antitrust laws are designed to protect."[8] The latter is an example of activities that are underperformed in the baseline economy and taxed by the policy intervention. Government tax and spending policies clearly involve some degree of classification error, especially when the policies feature bright-line proxies that trigger policy action, as with the 50-employee definition of a "large business" to be subject to additional

---

[7] Masoudi (2006) and Alden (2014) are two antitrust examples; see Harris (1970) on crime and punishment generally. Henderson (2006) discusses foreign-policy classification errors. See also the literature on the relationship between labor unions and income equality (Card et al., 2004 and the references cited therein).
[8] The two quotes are from Spectrum Sports, Inc. v. McQuillan (1993, para. 459) and Verizon Communications Inc. v. Law Offices of Curtis v. Trinko, LLP (2004, p. 3), respectively. See also the references cited in Devlin and Jacobs (2010).

regulation or a $23,850 annual income (i.e., the poverty line for a family of four) that indicates the withdrawal of social assistance.

Hayek, and more recently Manski (2005), emphasize that scientific knowledge is limited for policymaking purposes.[9] These limits are a source of the classification errors. But adequate knowledge in the minds of the scientists is not sufficient to eliminate such errors. Under a "political feasibility" interpretation, politically powerful groups have a particular interest in activities in the range $(T,T+m/2]$ so they induce the policymaker to give them the subsidy δ and to levy the corresponding tax on the politically weak who have a particular interest in activities in the interval $(T-m/2,T]$. Under a "political psychology" interpretation, scientific knowledge is not fully reflected in policy decisions because individuals participating in the polity have little incentive to fully process the scientific information that might improve public policy (Glaeser 2004, p. 410).

Note that I assume that all policies, even the imprecise ones, reflect the actual baseline frequency T of underperformed activities. By setting $\delta \leq \tau$, I also assume that policymakers have not exaggerated the magnitude of the baseline distortions. Finally, aside from the misclassification error, the imprecise policies are also assumed to be accurate in terms of the distribution of the baseline distortion: namely that the distortion is homogeneous among activities $[0,T]$. Relaxing any one of these assumptions would further strengthen this paper's conclusion that imprecise policies add to baseline productivity losses even when they share other qualitative features with the precise policy.

When the distortions satisfy (7) and (8), the cross-activity correlation ρ between the policy intervention $\ln(p'/p)$ and the baseline distortions $\ln p$ is shown in equation (9):

$$\rho = \frac{2m}{4(1-T)T} - 1 \tag{9}$$

The precise policy has $\rho = -1$ (zero misclassification). The correlation is zero when $2(1-T)T$ of the activities are misclassified.

---

[9] Interestingly, Friedman (1953) was published only eight years after Hayek published *The Use of Knowledge in Society* (1945) and three years after Hayek came to Chicago (although not in the same department as Friedman).

Because the prices and quantities are constant within each of the four segments noted in equation (8), log productivity is:[10]

$$\ln v = \frac{\theta}{\theta-1} \ln\left[\left(T - \frac{m}{2}\right) e^{(1-\theta)(\tau-\delta)} + \left(\frac{m}{2}\right) e^{(1-\theta)\tau} + \left(\frac{m}{2}\right) e^{-(1-\theta)\delta} + 1 - T - \frac{m}{2}\right]$$
$$- \ln\left[\left(T - \frac{m}{2}\right) e^{-\theta(\tau-\delta)} + \left(\frac{m}{2}\right) e^{-\theta\tau} + \left(\frac{m}{2}\right) e^{\theta\delta} + 1 - T - \frac{m}{2}\right] \tag{10}$$

B. A Negative Correlation is Insufficient for Tuned Interventions

Equation (10) is the basis for the main results in this section, some of which pertain to "tuned" policy interventions that have $\delta = \tau$ in order to fully offset at least some of the baseline distortions. A number of policy interventions are designed this way. The federal Earned Income Tax Credit is phased in (with annual income) at a 7.65 percent rate to offset employee federal payroll taxes, which are levied at the exact same rate.[11] Treble antitrust damages are often interpreted as an approximation of the expected amount of harm to customers from overcharges.[12] The magnitude of Affordable Care Act's employer-provided health insurance excise tax rate was set to offset the longstanding bias against individual-market insurance in favor of employer-provided insurance.[13]

The impact of tuned policy interventions on log productivity is summarized in equation (11):

$$\Delta \ln v = G(e^{-\tau}, T, \theta) - L(e^{-\tau}, m, \theta) \tag{11}$$

$$G(e^{-\tau}, T, \theta) \equiv \theta\left\{\ln\left[Te^{-\theta\tau} + 1 - T\right]^{1/\theta} - \ln\left[Te^{-(\theta-1)\tau} + 1 - T\right]^{1/(\theta-1)}\right\} \geq 0 \tag{12}$$

---

[10] The special case of $\theta = 1$ is of interest, in which case productivity is:
$$\ln v = -(\tau - \delta)T - \ln\left[\left(T - \frac{m}{2}\right) e^{-(\tau-\delta)} + \left(\frac{m}{2}\right) e^{-\tau} + \left(\frac{m}{2}\right) e^{\delta} + 1 - T - \frac{m}{2}\right]$$
[11] It has also been proposed to expand the EITC to offset employer payroll taxes (Marr and Huang 2014).
[12] Carlton and Perloff (2000, p. 603). But see Baker (1988), who argues that the amount of customer harm increases with the expected amount of anti-trust damages.
[13] See Gruber (2011) and Tahiro (2014).

$$L(e^{-\tau}, m, \theta) \equiv \theta \left\{ \ln\left[\left(\frac{m}{2}\right)\left(e^{-\theta\tau} + e^{\theta\tau}\right) + 1 - m\right]^{1/\theta} \right.$$

$$\left. - \ln\left[\left(\frac{m}{2}\right)\left(e^{-(\theta-1)\tau} + e^{(\theta-1)\tau}\right) + 1 - m\right]^{1/(\theta-1)} \right\} \geq 0 \tag{13}$$

Note that the net gain is additively separable in m and T. The T-only terms are the net gain G from a precise intervention and the m terms are the loss L from policy imprecision.

<u>Proposition 3</u> If $\delta = \tau$, $\theta \geq \frac{1}{2}$, and $T \leq \frac{1}{2}$, then, regardless of the magnitude of $\tau > 0$, a correlation of $\rho < -\frac{1}{2}$ is necessary for the policy intervention to increase productivity above the baseline.

The proof of Proposition 3 is in the appendix, which also shows exact values for the correlation required for a policy to enhance productivity. $T \leq \frac{1}{2}$ is a restriction that prevents consideration of a number of special cases. For $\theta \in (0,\frac{1}{2})$, the threshold correlation is closer to zero, although typically still near $-\frac{1}{2}$. This section focuses on $\theta \geq \frac{1}{2}$ because of its practical interest (that is, many applications feature activities that are not strong complements) and its similarity to Friedman's result.

The elasticity of substitution is an important determinant of the baseline productivity losses relative to no distortions. But Propositions 3 compares two distorted economies, so the question is whether it is more costly to (a) distort activities $[0,T]$ by τ relative to the others, as in the baseline, or (b) distort activities $[T-m/2, T+m/2]$ by τ relative to the others *and* 2τ within that interval. For small tax rates, the elasticity of substitution is irrelevant for this comparison.

What matters is the essentially quadratic nature of the productivity costs of the distortions, and the symmetry (or lack thereof) of the relationship between productivity and the various activities. Take, for example, $T = \frac{1}{2}$, which means that exactly half of the activities are overperformed in the baseline economy and the other half underperformed. If $m = \frac{1}{4}$ (i.e., $\rho = -\frac{1}{2}$) then that many activities are about twice as far from their efficient level as they are in the baseline,[14] which means that total quadratic deadweight costs are about the same as they are in the baseline.

---

[14] The remaining activities are at about the efficient level.

C. The quantity test is equivalent to the correlation test

The baseline allocation has too much of each of the activities $t \in (T, 1]$. Let X_1 denote the average amount of those activities (hereafter, "the average overperformed activity"). The amount X_1 is calculated as:

$$X_1 \equiv (1-T)^{-1} \int_T^1 x_t dt = (1-T)^{-1} \frac{\frac{m}{2} e^{\theta \delta} + 1 - T - \frac{m}{2}}{\left(T - \frac{m}{2}\right) e^{-\theta(\tau - \delta)} + \frac{m}{2} e^{-\theta \tau} + \frac{m}{2} e^{\theta \delta} + 1 - T - \frac{m}{2}} \quad (14)$$

A precise policy with $\delta = \tau$ would reduce X_1 to 1, which is its efficient value. However, Proposition 4 shows that an imprecise policy may reduce X_1 but nonetheless reduce productivity below the baseline.

<u>Proposition 4</u> The effect on X_1 (the average activity among those that were overperformed in the baseline) of any sized policy intervention $\delta > 0$ has the same sign as ρ.

In words, a quantity test for the policy impact on the average overperformed activity is weak: it only indicates whether the policy distortions are negatively correlated with the baseline distortions.

D. Productivity-maximizing caution

Taking as given that the policy intervention properly classifies activities often enough that $\rho < 0$, there is some intervention δ in $[0, \tau]$ that maximizes productivity relative to all other feasible policy interventions. For small tax rates, the productivity-maximizing δ is closely approximated by:

$$\frac{\delta}{\tau} \approx -\rho \tag{15}$$

At the productivity-maximizing δ, $1-\rho^2$ of the baseline productivity losses remain. Appendix II shows that equation (15) can be a good approximation for larger tax rates as well but, for especially large tax rates, large substitution elasticities, or (especially) both, the productivity-maximizing δ is significantly closer to zero than indicated by equation (15).

In words, despite classification errors there is a policy that can increase productivity above the baseline, if the intervention is small relative to the baseline distortions. The productivity increase could be small compared to the productivity losses that come from classification errors. For example, at $\rho = -1/3$, the productivity-maximizing intervention of about $\delta = \tau/3$ eliminates only 1/9 of the baseline productivity losses. The productivity losses from misclassification can be measured as 4/9 of the baseline losses because (a) $\delta = \tau/3$ and (b) $\rho = -1$ would be eliminating 5/9 of the baseline losses.

E. The critical correlation with a nested CES

A nested CES is an alternative setup in which one elasticity $\theta_b > 0$ governs substitution "between" [0,T] and (T,1] while another $\theta_w > 0$ governs substitution "within" each interval. From this perspective, the baseline economy (7) is unique because it has only between-type missallocations, whereas the tuned policy intervention primarily features within-type missallocations. If there were more substitution within than between, then one might guess that the critical correlation is even further from zero than $-\frac{1}{2}$. Appendix II proves that the critical correlation for the nested CES is (assuming small baseline distortions, as in Proposition 2):

$$\frac{1 - \sqrt{1 - (\theta_w/\theta_b) + (\theta_w/\theta_b)^2}}{(\theta_w/\theta_b) - 1} \in (-1,0) \tag{16}$$

which is a negative correlation. It is further from zero than $-\frac{1}{2}$ if and only if $\theta_w > \theta_b$ and has a limit of $-\frac{1}{2}$ as θ_w approaches θ_b, which is the case modeled by the production function (2). For

example, if the within substitution is twice as much as the between substitution, the critical correlation is $1 - \sqrt{3}$, which is about −0.73. In this case, in order for a tuned policy intervention to increase productivity relative to the baseline its taxes must be more negatively correlated with the baseline distortions than −0.73. If instead there were substitution within ($\theta_w > 0$) but no substitution between (θ_b approaching 0), then the critical correlation would approach −1 because the baseline economy is already free from within-type distortions.[15]

Take Starbucks, which has thousands of coffee stores most of which are company owned. With thousands of employees as part of its organization, Starbucks coexists in many markets with independent coffee shops and with franchised coffee shops like Dunkin Donuts.[16] Many of the organizations competing with Starbucks have fewer than 50 employees and are therefore subject to different taxes and regulations. Any policy that "accidentally" drives a wedge between Starbucks and its smaller competitors in the process of pursuing some other policy goal may be well described by the nested CES model with a large θ_w compared to θ_b.[17]

IV.   Conclusions

Cogley, et al. (2008), Friedman (1953), Hansen and Sargent (2007), and Manksi (2005) all make a case for policy analysis that explicitly incorporates imperfect knowledge about the policy effects. This paper complements that work in a couple of ways. First, as in some of the law and economic literature, my emphasis is with imperfections with the policy implementation, are not necessarily gaps in "scientific" knowledge about how people and businesses react to public policy.[18] I treat the elasticity of substitution θ, the size τ of the benchmark distortions,

---

[15] Total distortion variance is the sum of between variance and within variance, but in the nested-CES case the two sources of variance make different contributions to the productivity loss (the variance associated with more substitution counts more).

[16] Jargon (2013) and Dunkin' Donuts (2014) describe ownership structures for Starbucks and Dunkin' Donuts, respectively.

[17] Here I assume that Starbucks' production is a very close substitute for its competitors'.

[18] Tax policy design studies rarely account for the misallocations resulting from policy implementation errors, even though tax policy evaluation studies like Ballard, Shoven and Whalley (1985) and Gravelle (1989) have emphasized the costs of tax-induced misallocations.

and the scope *T* of the benchmark distortions, as known parameters. The problem is that the policymaker does not, or cannot, classify specific activities with 100 percent accuracy.

Second, the model is explicit about market distortions, incentives, and their relations with public policy. In the model, there is no such thing as an isolated policy "treatment" because a tax on one activity amounts to a subsidy on all other activities. The market distortions are associated with convex, and nearly quadratic, costs for the usual incentive reasons. Convexity is the fundamental reason why it is "difficult" (defined rigorously in this paper) for public policy to offset baseline distortions rather than add to them.

Hansen and Sargent have looked at policy rules in environments in which agents are averse to risk or uncertainty. Their 2014 paper finds that optimal policy may be more aggressive in the presence of such aversion than it would be otherwise, depending on how the policy instruments interact with uncertainty (Hansen and Sargent 2014). My model has no risk aversion or uncertainty, but emphasizes the "market's aversion" to deviations from the law of one price. These deviations have convex costs that result in "cautionary" policy rules. In my setting with policy implementation errors, that means that policy maximizes productivity by only partly offsetting baseline distortions. A policy that attempts to fully offset the baseline distortions results in less productivity if the correlation between the policy intervention and the baseline distortions is greater than –½.

Policy affects productivity by increasing the quantity of some activities while decreasing others. But a policy that increases the activities underperformed in the baseline economy may still result in less productivity to the extent that it distorts the mix of the once-underperformed activities. To put it another way, assessing qualitative features of a policy intervention such as the general direction of price and quantity impacts is not enough to determine even the direction of its productivity effect.

My model has no explicit rents and focuses on the misallocation costs, which are zero to first order and convex, of misclassifying activities for policy purposes. As a result, sufficiently cautious policy interventions are an opportunity to increase productivity above the baseline, at least if the interventions are precise enough to pass the quantity test. But misclassification presumably encourages rent-seeking behavior as those who would be taxed by the precise

corrective policy seek to be subsidized instead, or at least get relief from the corrective taxes. Take, for example, a Blanchard and Kiyotaki (1987) setup where just a few "owner" agents have the right to engage in any one activity t, and thereby enjoy a degree of market power. As a result, a fraction of aggregate income accrues as a rent on those rights. Policy interventions reallocate those rents, and the owners may rationally expend real resources trying to influence how they are classified. Depending on the details of the rent-seeking contest (Tullock 1998), the resources used in rent seeking could be a first-order cost and thereby broaden the conditions under which the baseline economy achieves the maximum productivity despite its distortions.

V. Appendix I: Proof of Propositions 1 and 3

Proposition 1 Lognormally distributed market distortions create log productivity losses that are proportional to the elasticity of substitution and cross-activity variance of the price profile.

Proof   Substituting the equilibrium quantities (equation (3)) into the production function (2),

$$\ln v = \frac{\theta}{\theta - 1} \ln \left[\int_0^1 e^{(1-\theta) \ln p_t} dt \right] - \ln \left[\int_0^1 e^{-\theta \ln p_t} dt \right] \tag{17}$$

The lognormal distribution has

$$\ln p_t = \mu + \sigma \Phi^{-1}(t) \tag{18}$$

where Φ^{-1} is the inverse standard normal distribution, $\sigma > 0$ is the cross-activity standard deviation of $\ln p$, and μ is the mean. Substitution of (18) into (17) and integrating yields

$$\ln v = \frac{\theta}{\theta - 1} \left[\frac{\sigma^2}{2} (\theta - 1)^2 - (\theta - 1)\mu \right] - \left[\frac{\sigma^2}{2} \theta^2 - \theta\mu \right] = -\frac{\sigma^2}{2} [\theta^2 - (\theta - 1)\theta]$$
$$= -\frac{\theta \sigma^2}{2} \tag{19}$$

Proposition 3 If $\delta = \tau$, $\theta \geq \frac{1}{2}$, and $T \leq \frac{1}{2}$, then, regardless of the magnitude of $\tau > 0$, a correlation of $\rho < -\frac{1}{2}$ is necessary for the policy intervention to increase productivity above the baseline.

Proof   The intervention's impact on log productivity is shown in equation (11). As a function of the correlation ρ, the impact is:

$$\Delta \ln v = G(e^{-\tau}, T, \theta) - L(e^{-\tau}, (1+\rho)2(1-T)T, \theta) \tag{20}$$

with the functions G and L defined in the main text. The proof has three versions: graphical, numerical, and (for the special case $\theta = 1$ and the special case $\tau \to \infty$) analytical. The level surfaces of zero impact are, for selected values of θ, plotted below on the entire domain cited in the proposition: $\rho \in [-1, -\frac{1}{2}]$, $\tau > 0$, $T \in [0, \frac{1}{2}]$. For a given θ, the set below a level surface contains the parameters yielding net gains. Note that all of the level surfaces are below $\rho = -\frac{1}{2}$.

[Figures 2 and 3 here]

For a numerical demonstration, I randomly sampled 10 million parameter vectors satisfying $e^{-\tau} \in (0,1)$, $T \in (0,\frac{1}{2})$, $\theta \in (\frac{1}{2},\infty)$, and $\rho \in (-\frac{1}{2},1)$. None of them showed a net gain.

Equations (21)-(28) below are an analytical proof for $\theta = 1$. The policy impact is:

$$\Delta \ln v = T\tau + \ln[Te^{-\tau} + 1 - T] - \ln[1 + (1 + \rho)2(1 - T)T(\cosh \tau - 1)] \tag{21}$$

where cosh denotes the hyperbolic cosine function. For $T \in (0,1)$ and $\tau > 0$, the net gain is strictly decreasing in ρ. Defining $z = e^{-\tau}$, the correlation that results in a zero net gain is:

$$\rho(z,T) = \frac{(Tz + 1 - T)z^{-T} - 1}{(z + z^{-1} - 2)(1 - T)T} - 1 \tag{22}$$

At $T = \frac{1}{2}$, the threshold correlation is:

$$\rho\left(z,\frac{1}{2}\right) = 2\frac{z^{\frac{1}{2}} + z^{-\frac{1}{2}} - 2}{z + z^{-1} - 2} - 1 \leq -\frac{1}{2} \tag{23}$$

The inequality follows from the facts that (a) $z \in [0,1]$, (b) $\rho(1,\frac{1}{2}) = -\frac{1}{2}$, and (c) the ratio in equation (23) is weakly increasing in z. Fact (c) comes from (a) and from differentiating (23) with respect to z.

Lemma 1 says that $\rho(z,T) \leq \rho(z,\frac{1}{2})$ for $T \leq \frac{1}{2}$, so it and (23) prove the Proposition for $\theta = 1$.

Lemma 1 $\rho(z,T)$ is weakly increasing with T on $T \in [0,\frac{1}{2}]$.

Proof $\rho_T(z,T)$ is:

$$\rho_T(z,T) = z^{1-T}\frac{(z^T - 1)(1 - 2T) + (z - 1)T^2 - [1 - (1 - z)T](1 - T)T \ln z}{(1 - z)^2(1 - T)^2T^2} \tag{24}$$

The sign of $\rho_T(z,T)$ is therefore the sign of the numerator $f(z,T)$, which has a Taylor-series representation:

$$f(z,T) = -T^2(1-z) - \sum_{n=1}^{\infty} A_n(T)(1-z)^n$$
$$+ [1 - (1-z)T](1-T)T \sum_{n=1}^{\infty} B_n(1-z)^n \tag{25}$$

where the first sum represents the $(z^T{-}1)$ term and the second sum represents the $\ln z$ term. Note that equation (25) is not an approximation because the sum goes to infinity. The coefficients are (for integers $n > 1$):

$$A_n(T) = \frac{n-1-T}{n} A_{n-1}(T), \qquad A_1(T) = (1 - 2T)T$$
$$B_n = \frac{n-1}{n} B_{n-1} = \frac{1}{n}, \qquad B_1 = 1 \tag{26}$$

Note that each A and each B are positive, unless $T = \frac{1}{2}$ in which case each A is zero. Collecting powers of $(1\text{-}z)$, the $n = 1$ and $n = 2$ terms drops out of the numerator function and

$$f(z,T) = \sum_{n=3}^{\infty} [R_n(T) - 1]A_n(T)(1-z)^n \tag{27}$$

$$R_n(T) = \frac{(1-T)n - 1}{(1-T)(n-1) - 1} \frac{n-2}{n-1-T} R_{n-1}(T) > 1, \qquad R_2(T) = 1 \tag{28}$$

Equations (27) and (28) prove that the numerator $f(z,T)$ of $\rho_T(z,T)$ is nonnegative, because $f(z,T)$ is the sum of a product with each element of the product being nonnegative on the domain $z \in [0,1]$ and $T \in [0, \frac{1}{2}]$.

With $\theta > 1$, the limiting case $\tau \to \infty$ is instructive because it shows that values of ρ above $-\frac{1}{2}$ result in productivity losses but those losses are not increasing with ρ.[19] In this limit, productivity is positive but only the activities $(T,1]$ are performed in the baseline economy, and only the activities $(T,T{+}m/2)$ are performed as a result of the policy intervention. The productivity impact is:

$$\Delta \ln v = \frac{1}{\theta - 1} \ln(1 + \rho)T \leq 0 \tag{29}$$

The equality is strict if $T < \frac{1}{2}$. The impact is zero with $T = \frac{1}{2}$ and $\rho = 1$ because the policy intervention just swaps the equal-sized intervals of tasks that are not performed.

---

[19] With $\theta \leq 1$, productivity goes to zero as $\tau \to \infty$ and the policy intervention can improve upon the baseline only if there is zero classification error ($\rho = -1$).

VI. Appendix II: Taylor approximations used in the Paper

The end of the paper uses approximation methods to (a) approximate the optimal degree of policy "caution" and (b) show the critical correlation for a nested CES model in the neighborhood of small baseline tax rates. Below are more details on those approximations.

A second-order approximation of productivity is used to calculate the optimal degree of caution (see equation (15)). Table 1 below provides two indicators of the quality of this approximation for selected substitution elasticities θ, as estimated numerically by evenly gridding the parameter space $[\tau, T, \rho]$, but excluding $\tau > 1$.[20] The exact solution for δ/τ was, on average, closer to zero than the approximate solution (15). The standard deviation of the approximation error, which has units of a correlation, is no greater than 0.1 as long as the elasticity of substitution is no greater than two.

[Table 1 here]

The nested CES model distinguishes activities in the interval $[0,T]$ from those in $(T,1]$. The substitution elasticity within either interval is θ_w. The substitution elasticity between an activity that is in the first interval and an activity in the second interval is θ_b. Productivity is:

$$v = CES\left(CES\left(x_1, x_2, 1 - \frac{m}{2T}, \theta_w\right), CES\left(y_2, y_1, 1 - \frac{m}{2(1-T)}, \theta_w\right), T, \theta_b\right) \tag{30}$$

$$CES(x_1, x_2, \alpha, \theta) \equiv \left[\alpha x_1^{(\theta-1)/\theta} + (1-\alpha)x_2^{(\theta-1)/\theta}\right]^{\theta/(\theta-1)} \tag{31}$$

where x_1, x_2, y_1, y_2 denote the average quantities in the four intervals specified in the policy misspecification model (8), respectively. The baseline economy does not have within-interval distortions. In this case, the quantities satisfy $x_1 = x_2 < y_1 = y_2$ and are the same as in equation (3) with $\theta = \theta_b$. For the same reason, the baseline economy has the same productivity in nested CES model as in the single-CES model of the main text.

---

[20] τ was gridded from 0.1 to 1 in steps of 0.1, and thereby excluded the most prohibitive tax rates for which the approximation (15) is poor (the optimal δ is much closer to zero than the approximation). T was gridded from 0.01 to 0.99 in steps of 0.01. The correlation ρ was gridded from −0.95 to −0.05 in steps of 0.05. This makes a total of 18,810 grid points for each value of θ.

As in the main text (recall Figure 1), I denote $z = e^{-\tau} \leq 1$ as the price, with the policy intervention, of y_2 relative to x_2 and the price of y_1 relative to y_2. The equilibrium quantities and productivity satisfy:

$$\frac{v}{x_1} = CES\left(1, z^{\theta_w}, 1 - \frac{m}{2T}, \theta_w\right) CES\left(1, \left[\frac{CES\left(1, z^{-\theta_w}, 1 - \frac{m}{2(1-T)}, \theta_w\right)}{CES\left(1, z^{\theta_w}, 1 - \frac{m}{2T}, \theta_w\right)}\right]^{\theta_b/\theta_w}, T, \theta_b\right) \quad (32)$$

$$x_1^{-1} = TCES\left(1, z^{\frac{\theta_w^2}{\theta_w-1}}, 1 - \frac{m}{2T}, \theta_w\right)^{\frac{\theta_w-1}{\theta_w}} +$$

$$(1-T)CES\left(1, z^{\frac{\theta_w^2}{1-\theta_w}}, 1 - \frac{m}{2(1-T)}, \theta_w\right)^{\frac{\theta_w-1}{\theta_w}} \left[\frac{CES\left(1, z^{-\theta_w}, 1 - \frac{m}{2(1-T)}, \theta_w\right)}{CES\left(1, z^{\theta_w}, 1 - \frac{m}{2T}, \theta_w\right)}\right]^{\theta_b/\theta_w - 1} \quad (33)$$

$$m = (1+\rho)2(1-T)T \quad (34)$$

The critical correlation is the value of ρ such v is equal to the baseline productivity. As in the single-CES model in the main text, the critical correlation varies with z and T. In the limit as z goes to zero, the critical correlation is the formula (16) shown in the main text.

Table 1. Exact and approximate estimates of the productivity-maximizing policy intervention
Selected values of the elasticity of substitution

Elasticity of substitution	Exact δ/τ minus approximate	Root mean-squared deviation
0.5	-0.001	less than 0.01
1	-0.006	0.03
2	-0.036	0.10
5	-0.203	0.30

Notes : A second-order approximation of productivity implies that it is maximized at $\delta/\tau = -\rho.$ Each row above is calculated with 18,810 evenly spaced grid points: τ from 0.1 to 1 in intervals of 0.1, T from 0.01 to 0.99 in intervals of 0.01, and ρ from -0.95 to -0.05 in intervals of 0.05.

Figure 1. Price profiles with and without policy intervention.

m > 0 denotes activities that are misclassified. *t* is an index of activities.
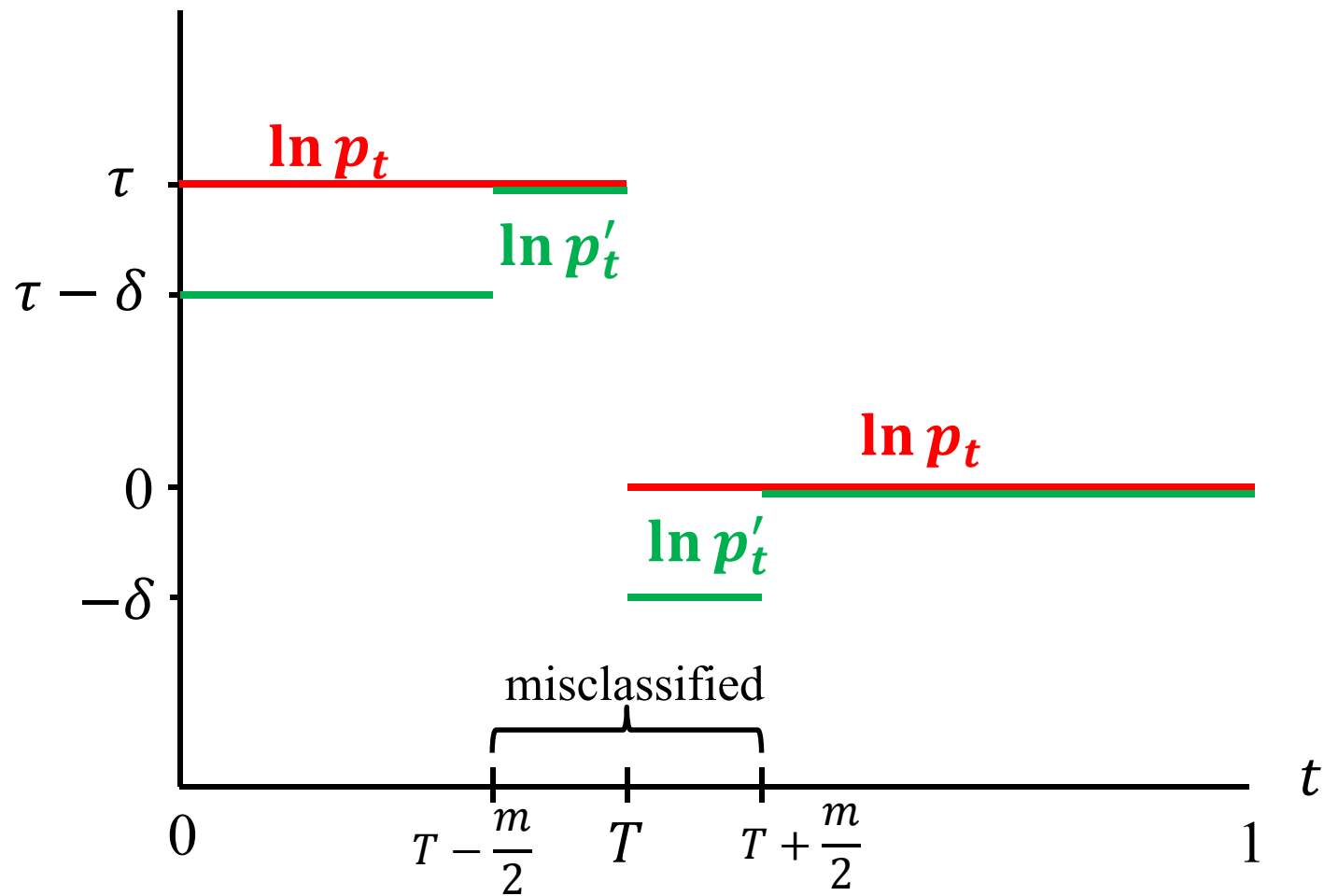
Figure 2. Zero-net-gain contours for selected elasticities of substitution

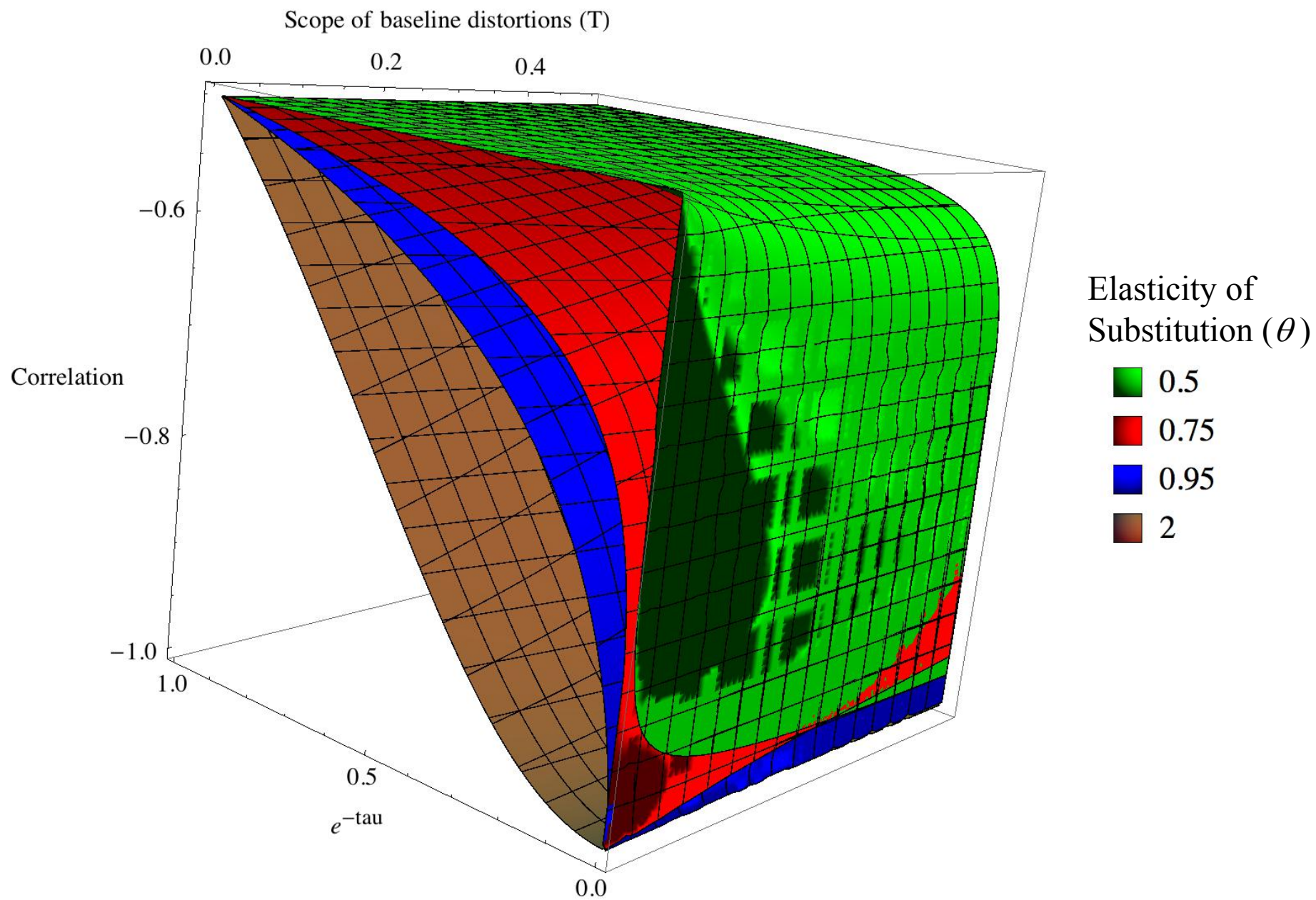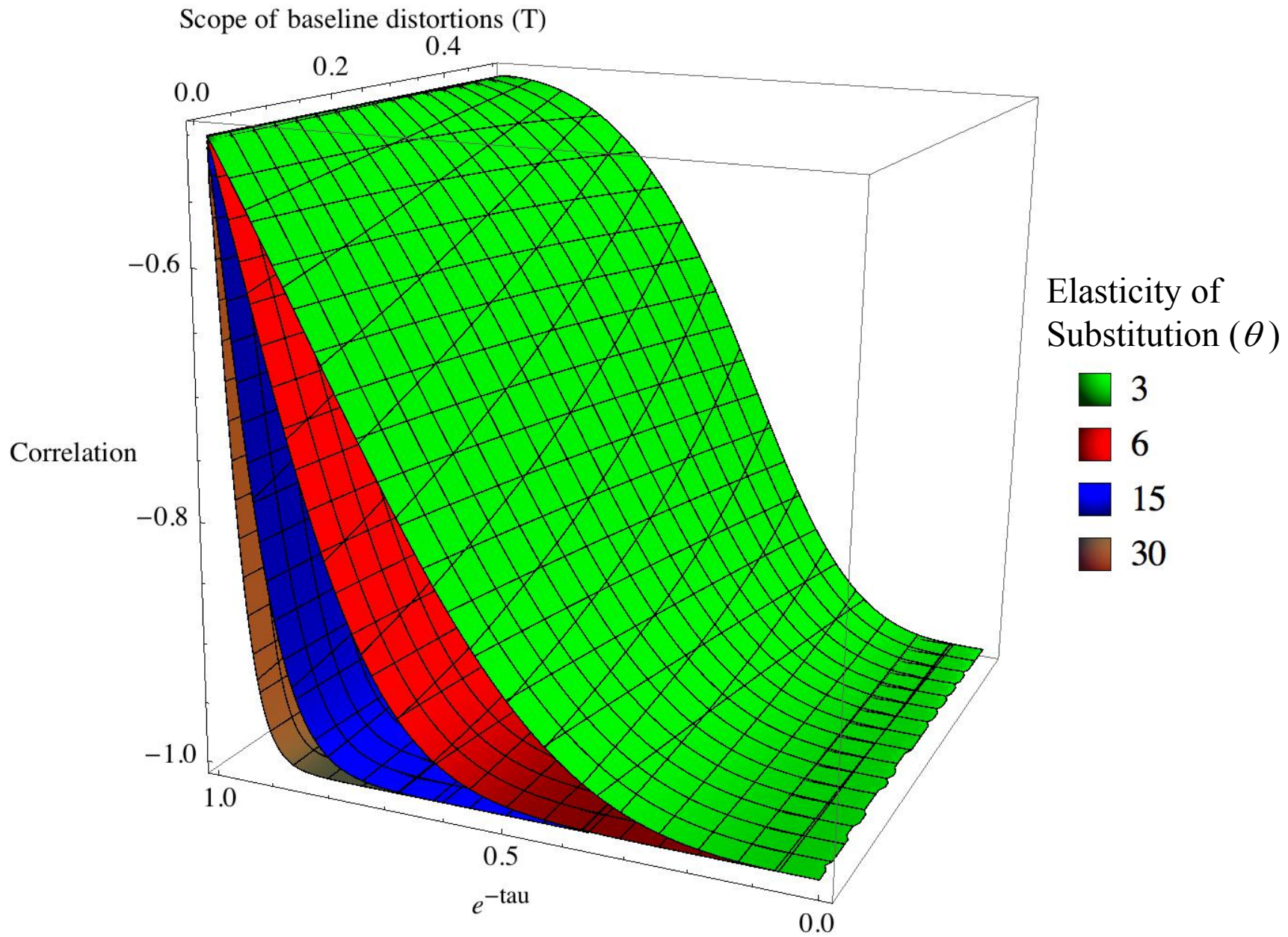Policy intervention tuned to baseline distortions ($\delta = \tau$).

Figure 3. Zero-net-gain contours for selected elasticities of substitution
Policy intervention tuned to baseline distortions ($\delta = \tau$).

Bibliography

Alden, Abbott. "A Cost-Benefit Framework for Antitrust Enforcement Policy." *The Heritage Foundation.* July 17, 2014. http://www.heritage.org/research/commentary/2014/7/a-cost-benefit-framework-for-antitrust-enforcement-policy (accessed December 31, 2014).

Baker, Jonathan B. "Private Information and the Deterrent Effect of Antitrust Damage Remedies." *Journal of Law, Economics & Organization* 4, no. 2 (Autumn 1988): 385-408.

Ballard, Charles L., John B. Shoven, and John Whalley. "General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States." *American Economic Review* 75, no. 1 (1985): 128-38.

Blanchard, Olivier Jean, and Nobuhiro Kiyotaki. "Monopolistic Competition and the Effects of Aggregate Demand." *American Economic Review* 77, no. 4 (September 1987): 647-66.

Card, David, Thomas Lemieux, and W. Craig Riddell. "Unions and Wage Inequality." *Journal of Labor Research* 25, no. 4 (Fall 2004): 519-59.

Carlton, Dennis W., and Jeffrey M. Perloff. *Modern Industrial Organization.* 3rd edition. New York: Addison Wesley Longman, 2000.

Chetty, Raj. "When do Countercyclical Policies Increase Economic Stability?" *rajchetty.com.* September 2001. http://obs.rc.fas.harvard.edu/chetty/files/stabilization.pdf (accessed November 1, 2014).

Cogley, Timothy, Ricardo Colacito, Lars Peter Hansen, and Thomas J. Sargent. "Robustness and U.S. Monetary Policy Experimentation." *Journal of Money, Credit, and Banking* 40, no. 8 (December 2008): 1599-1623.

Devlin, Alan, and Michael Jacobs. "Antitrust Error." *William and Mary Law Review* 52, no. 1 (2010): Article 3.

Friedman, Milton. "The Effects of Full Employment Policy on Economic Stability: A Formal Analysis." In *Essays in Positive Economics*, by Milton, ed. Friedman, 117-32. Chicago: University of Chicago Press, 1953.

Glaeser, Edward L. "Psychology and the Market." *American Economic Review* 94, no. 2 (May 2004): 408-13.

Gravelle, Jane G. "Differential Taxation of Capital Income: Another Look at the 1986 Tax Reform Act." *National Tax Journal* 42, no. 4 (December 1989): 441-63.

Gruber, Jonathan. "The Tax Exclusion for Employer-sponsored Health Insurance." *National Tax Journal* 64, no. 2, Part 2 (June 2011): 511-30.

Hansen, Lars Peter, and Thomas J. Sargent. *Four Types of Ignorance.* manuscript, New York University, 2014.

—. *Robustness.* Princeton, NJ: Princeton University Press, 2007.

Harberger, Arnold C. "The Incidence of the Corporation Income Tax." *Journal of Political Economy* 70, no. 3 (June 1962): 216-40.

Harris, John R. "On the Economics of Law and Order." *Journal of Political Economy* 78, no. 1 (January 1970): 165-74.

Hayek, Friedrich A. "The Use of Knowledge in Society." *American Economic Review* 35, no. 4 (September 1945): 519-30.

Henderson, David R. "Milton and David Friedman on Military Intervention." *antiwar.com.* January 9, 2006. http://www.antiwar.com/henderson/?articleid=8347 (accessed December 31, 2014).

Hsieh, Chang-Tai, and Peter J. Klenow. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124, no. 4 (November 2009): 1403-48.

Manski, Charles F. *Social Choice with Partial Knowledge of Treatment Response.* Princeton, NJ: Princeton University Press, 2005.

Marr, Chuck, and Chye-Ching Huang. "Strengthening the EITC for Childless Workers Would Promote Work and Reduce Poverty." *Center for Budget and Policy Priorities.* July 30, 2014. http://www.cbpp.org/cms/?fa=view&id=3991 (accessed December 31, 2014).

Masoudi, Gerald F. *The Best Approach to Enforcement Against Single-Firm Conduct: Caution.* Presentation, Washington, DC: ABA Antitrust Section's Fall Forum, 2006.

Merton, Robert C. "Lifetime Portfolio Selection under Uncertainty: The Continuous-time Case." *Review of Economics and Statistics* 51, no. 3 (August 1969): 247-57.

Restuccia, Diego, and Richard Rogerson. "Policy Distortions and Aggregate Productivity." *Review of Economic Dynamics* 11, no. 4 (October 2008): 707-20.

Rosen, Sherwin. "The Theory of Equalizing Differences." In *Handbook of Labor Economics, vol. I*, by Orley C. Ashenfelter and Richard, eds. Layard, 641-92. Amsterdam: North-holland, 1986.

Shoven, John B. "The Incidence and Efficiency Effects of Taxes on Income from Capital." *Journal of Political Economy* 84, no. 6 (December 1976): 1261-83.

Spectrum Sports, Inc. v. McQuillan. 506 US 447 (U.S. Supreme Court, January 25, 1993).

Tahiro, Jordan. "The Obamacare Cadillac Tax and its Mixed Bag of Consequences." *The Tax Policy Blog.* June 12, 2014. http://taxfoundation.org/blog/obamacare-cadillac-tax-and-its-mixed-bag-consequences (accessed December 31, 2014).

Tullock, Gordon. "The Fundamentals of Rent-Seeking." *The Lock Luminary* I, no. 2 (Winter 1998).

Verizon Communications Inc. v. Law Offices of Curtis v. Trinko, LLP. 540 US 398 (U.S. Supreme Court, January 13, 2004).