ONLINE APPENDICES FOR “Bargaining and Dynamic Competition” by Deng, Jia, Leccese and Sweeting

April 2024

A Analytical Proofs

A.1 Proofs of Proposition 1.

Recall Proposition 1.

Proposition 1 *In a model with any $m \leq M$,*

1. *if $\tau = 1$, equilibrium prices will equal marginal production costs in all states, for all ρ and δ.*
2. *there will be a unique symmetric MPNE when*
   (a) $\delta = 0$ for all ρ and τ, or
   (b) $\tau = 1$ for all ρ and δ.

Proof of Part 1. The structure of the proof is to show that VS must be zero in every state, which implies that price will equal marginal production costs. From text equation (9), $\tau = 1$ implies that all markups Φ are equal to zero. Equation (11) therefore implies that $\mathbf{VS_1} = 0$, and equation (10) implies $\widehat{\mathbf{c_1}} = \mathbf{c_1}$, so that equation (9) now implies $p_1^*(\mathbf{e}) = c_1(\mathbf{e})$ in all states.

Proof of Part 2(a). Follows from the recursive proof of Besanko, Doraszelski, Kryukov, and Satterthwaite (2010), and the fact logit demand implies there can only be one (static) MPNE in the absorbing state (M, M).

Proof of Part 2(b). Uniqueness of prices follows immediately from the proof of part 1, and the choice probabilities of a static seller are unique given prices.

A.2 Proof of Proposition 2.

Proposition 2 *For $m = M = 2$ and $\delta = 0$, the unique symmetric equilibrium will have the following properties*

1. *equilibrium $D_1^*(1,2)(\tau) < \frac{1}{2}$ for all τ.*
2. *equilibrium $D_1^*(1,2)(\tau)$ is strictly decreasing in τ.*
3. *for $t \geq 2$, HHI^t is strictly increasing in τ.*
4. *there exists a τ^* such that $TS^{PDV}(\tau^*) = TS^{SP}$, $TS^{PDV}(\tau)$ is strictly increasing in τ for $\tau \in (0, \tau^*)$ and strictly decreasing in τ for $\tau \in (\tau^*, 1)$.*

A.2.1 Characterization

Given $M = 2$ and $\delta = 0$, the equilibrium is fully characterized by the equilibrium condition for $D_1^*(1,2)$, which can be simplified to

$$H(x,\tau) - H\left(\frac{1}{2}, \tau\right) = c(1) - c(2), \tag{18}$$

where $H(x)$ is defined as

$$H(x,\tau) := \log\frac{1-x}{x} - \frac{1-\beta+\beta x}{1-\beta}\left[\widetilde{\Phi}\left(x, \frac{\tau}{1-\tau}\right) - (1-\beta)\widetilde{\Phi}\left(1-x, \frac{\tau}{1-\tau}\right)\right] \text{ for } x, \tau \in [0,1],$$

where $\widetilde{\Phi}(x,z)$ is defined as

$$\widetilde{\Phi}(x,z) := \frac{\log\frac{1}{1-x}}{zx + (1-x)\log\frac{1}{1-x}} \text{ for } x \in [0,1] \text{ and } z \in [0,\infty),^{19}$$

[19] $\widetilde{\Phi}(x,\infty)$ is defined to be 0 for all $x \in [0,1]$.

and is a reformulated version of the mark-up condition $\Phi(x,\tau)$ where the second argument in $\widetilde{\Phi}(x,z)$ is $\frac{\tau}{1-\tau}$.

As negotiated prices are a transfer from buyers to sellers, we can express expected total surplus in any state as a function of the choice probabilities and firm costs only, i.e., $\sum_{k=1,2} D_k(\mathbf{e})\left(\log\frac{1}{D_k(\mathbf{e})} - c_k(\mathbf{e})\right)$ which does not depend on τ. In the $M=2$ and $\delta=0$ case, TS^{PDV}, for a game starting at (1,1), can be written as the following function of $D_1(1,2)$,

$$\begin{aligned} TS^{PDV}(x) =& \frac{\beta}{1-\beta+\beta x}\left[x\log\frac{1}{x} + (1-x)\log\frac{1}{1-x} - (c(1)-c(2))x - \log 2\right] + \\ & \frac{\log 2 - \beta c(2) - (1-\beta)c(1)}{1-\beta}. \end{aligned}$$

We define D_1^{SP} as the choice probability that the social planner would choose in state (1,2). The social planner, would, of course, use choice probabilities of $\frac{1}{2}$ in states (1,1) and (2,2) where the suppliers are symmetric, with the firm with the highest ε making the sale.

When $\delta=0$ and $x=D_1(1,2)$,

$$HHI^t = (x^2+(1-x)^2)(1-x)^{t-1} + \frac{1}{2}(1-(1-x)^{t-1})$$

for $t\geq 2$ as the state will either be (1,2) with concentration $(x^2+(1-x)^2)$ or (2,2) with concentration $\frac{1}{2}$.

A.2.2 Preliminary Results.

Lemma A.1 *$H(x,\tau)$ is strictly decreasing in $x\in(0,1)$.*

Proof of Lemma A.1. Since there is a unique equilibrium when $\delta=0$ for any parameterization of $c(e)$, $H(x,\tau)$ must be strictly monotone in $x\in(0,1)$. Otherwise, there will be multiple equilibria—i.e., multiple solutions to (18)—for some value of $c(1)-c(2)$.

It suffices to show that $H(x,\tau)$ is decreasing on some interval in $(0,1)$. When $\tau = 1$, it is easy to see that $H(x,\tau)$ is strictly decreasing $x \in (0,1)$. When $\tau < 1$, we have that $\widetilde{\Phi}\left(x, \frac{\tau}{1-\tau}\right)$ strictly increases with $x \in (0,1)$. Therefore, $H(x,\tau)$ decreases with x when $\widetilde{\Phi}\left(x, \frac{\tau}{1-\tau}\right) > (1-\beta)\widetilde{\Phi}\left(1-x, \frac{\tau}{1-\tau}\right)$. This completes the proof. ■

Lemma A.2 $TS^{PDV}(x)$ *is strictly increasing in* $x \in (0, D_1^{SP})$ *and strictly decreasing in* $x \in (D_1^{SP}, 1)$, *where* $D_1^{SP} \in (0, 1/2)$ *solves*

$$\log\frac{1}{x} - \frac{1}{1-\beta}\log\frac{1}{1-x} = c(1) - c(2) - \frac{\beta}{1-\beta}\log 2. \tag{19}$$

Proof of Lemma A.2. The proof immediately follows from the fact that

$$\frac{dTS^{PDV}(x)}{dx} = \frac{\beta(1-\beta)}{(1-\beta+\beta x)^2}\left\{\log\frac{1}{x} - \frac{1}{1-\beta}\log\frac{1}{1-x} - \left[c(1) - c(2) - \frac{\beta}{1-\beta}\log 2\right]\right\}.$$

■

A.2.3 Proofs of Proposition 2

Proof of Proposition 2.1: equilibrium $D_1^*(1,2)(\tau) < \frac{1}{2}$ for all τ. By Lemma A.1, the left-hand side of (18) is strictly decreasing in x. It is clear that the left-hand side equals 0 when $x = \frac{1}{2}$. Since the right-hand side is strictly positive, the solution must be less than $\frac{1}{2}$. ■

Proof of Proposition 2.2: equilibrium $D_1^*(1,2)(\tau)$ is strictly decreasing in τ. Applying the implicit function theorem to (18) yields that

$$\frac{\partial x}{\partial \tau} = -\frac{H_\tau(x,\tau) - H_\tau\left(\frac{1}{2},\tau\right)}{H_x(x,\tau)}.$$

By Lemma A.1, $H_x(x,\tau) < 0$, and by Proposition 2.1, $x < \frac{1}{2}$. So to complete the proof, it suffices to show that $H_\tau(x,\tau) < H_\tau\left(\frac{1}{2},\tau\right)$ for $x \in \left(0,\frac{1}{2}\right)$.

Note that

$$H_\tau(x,\tau) = -\frac{1-\beta+\beta x}{1-\beta}\left[\widetilde{\Phi}_z\left(x,z\right) - (1-\beta)\widetilde{\Phi}_z\left(1-x,z\right)\right]\frac{dz}{d\tau}, \text{ with } z = \frac{\tau}{1-\tau}.$$

Therefore,

$$\begin{aligned}
&H_\tau(x,\tau) < H_\tau\left(\frac{1}{2},\tau\right)\\
\iff\ & -\frac{1-\beta+\beta x}{1-\beta}\left[\widetilde{\Phi}_z\left(x,z\right) - (1-\beta)\widetilde{\Phi}_z\left(1-x,z\right)\right] < -\frac{(2-\beta)\beta}{2(1-\beta)}\widetilde{\Phi}_z\left(\frac{1}{2},z\right)\\
\iff\ & -\left[\widetilde{\Phi}_z\left(x,z\right) - (1-\beta)\widetilde{\Phi}_z\left(1-x,z\right)\right] < -\frac{(1-\beta)(2-\beta)\beta}{2(1-\beta)(1-\beta+\beta x)}\widetilde{\Phi}_z\left(\frac{1}{2},z\right).
\end{aligned} \tag{20}$$

In fact,

$$\widetilde{\Phi}_z(x,z) = -\frac{x\log\frac{1}{1-x}}{\left[zx+(1-x)\log\frac{1}{1-x}\right]^2}, \text{ and } \widetilde{\Phi}_z\left(\frac{1}{2},z\right) = -\frac{2\log 2}{(z+\log 2)^2} < 0.$$

So the right-hand side of (20) is decreasing in $x \in \left(0,\frac{1}{2}\right)$.

Next, we show that the left-hand side of (20) increases with $x \in \left(0,\frac{1}{2}\right)$. It suffices to show that $-\widetilde{\Phi}_z\left(x,z\right)$ increases with $x \in \left(0,\frac{1}{2}\right)$. In fact,

$$\begin{aligned}
-\widetilde{\Phi}_{zx}\left(x,z\right) &= \frac{\left(\log\frac{1}{1-x}+\frac{x}{1-x}\right)\left[zx+(1-x)\log\frac{1}{1-x}\right] - 2x\log\frac{1}{1-x}\left(z+1-\log\frac{1}{1-x}\right)}{\left[zx+(1-x)\log\frac{1}{1-x}\right]^3}\\
&= \frac{zx\left(\frac{x}{1-x}-\log\frac{1}{1-x}\right)+\log\frac{1}{1-x}\left[(1+x)\log\frac{1}{1-x}-x\right]}{\left[zx+(1-x)\log\frac{1}{1-x}\right]^3}.
\end{aligned}$$

It is easy to verify that $\frac{x}{1-x} > \log\frac{1}{1-x} > \frac{x}{1+x}$ for $x \in \left(0,\frac{1}{2}\right)$. As a result, $-\widetilde{\Phi}_{zx}\left(x,z\right) > 0$.

Combining the facts that the right-hand side of (20) is decreasing in $x \in \left(0,\frac{1}{2}\right)$, the left-hand side of (20) is increasing in $x \in \left(0,\frac{1}{2}\right)$, and both sides are equal at $x=\frac{1}{2}$, we can conclude that (20) holds for $x \in \left(0,\frac{1}{2}\right)$. This completes the proof. ■

Proof of Proposition 2.3: equilibrium expected concentration (HHI^t) in any period $t \geq 2$, strictly increases in τ. The derivative of HHI^t with respect to x is, for $t \geq 2$

$$-\frac{(1-2x)(1-x)^{t-2}(3+t-2x-2tx)}{2}$$

which is negative for any $x < \frac{1}{2}$. As equilibrium $D_1^*(1,2) < \frac{1}{2}$ (Proposition 2.1) and equilibrium $D_1^*(1,2)$ decreases in τ (Proposition 2.2), HHI^t increases in τ. ■

Proof of Proposition 2.4 : there exists a τ^* such that $TS^{PDV}(\tau^*) = TS^{SP}$, and $TS^{PDV}(\tau^*)$ is increasing in τ for $\tau \in (0, \tau^*)$ and decreasing in τ for $\tau \in (\tau^*, 1)$.

As TS^{PDV} only depends on $D_1^*(1,2)$, $TS^{PDV}(\tau^*) = TS^{SP}$ if $D_1^*(1,2)(\tau^*) = D_1^{SP}$. Because $D_1^*(1,2)(\tau)$ decreases with τ (Proposition 2.2), and Lemma A.2, the result follows if

$$D_1^*(1,2)(\tau = 0) > D_1^{SP} > D_1^*(1,2)(\tau = 1).$$

By (18) and the monotonicity of $H(x,\tau)$ in x (Lemma A.1), the above inequality is equivalent to

$$H(D_1^{SP},0)-H\left(\frac{1}{2},0\right)-[c(1)-c(2)] > 0 > H(D_1^{SP},1)-H\left(\frac{1}{2},1\right)-[c(1)-c(2)]. \quad (21)$$

We first show that $0 > H(D_1^{SP},1) - H\left(\frac{1}{2},1\right) - [c(1)-c(2)]$. In fact,

$$H(D_1^{SP},1) - H\left(\frac{1}{2},1\right) - [c(1)-c(2)] = \frac{\beta}{1-\beta}\left(\log\frac{1}{1-D_1^{SP}} - \log 2\right) < 0,$$

where the equality is due to that D_1^{SP} solves (19), and the inequality is due to that $D_1^{SP} < \frac{1}{2}$ (Lemma A.2).

It only remains to show that $H(D_1^{SP},0) - H\left(\frac{1}{2},0\right) - [c(1)-c(2)] > 0$. Note that

$$\begin{aligned}
&H(D_1^{SP},0) - H\left(\frac{1}{2},0\right) - [c(1)-c(2)] \\
&= \log\left(\frac{1-D_1^{SP}}{D_1^{SP}}\right) - [c(1)-c(2)] - \left[\frac{1}{(1-\beta)(1-D_1^{SP})} - \frac{1-\beta}{D_1^{SP}} - \frac{\beta(2-\beta)}{1-\beta}\right] - H\left(\frac{1}{2},0\right).
\end{aligned}$$

Again, substituting (19) into the right-hand side of the above equation yields that

$$H(D_1^{SP},0) - H\left(\frac{1}{2},0\right) - [c(1)-c(2)] = L(D_1^{SP}) - L\left(\frac{1}{2}\right),$$

where

$$L(x) := \frac{\beta}{1-\beta}\log\frac{1}{1-x} - \left[\frac{1}{(1-\beta)(1-x)} - \frac{1-\beta}{x}\right].$$

Since

$$L'(x) = -\frac{1-\beta(1-x)}{(1-\beta)(1-x)^2} - \frac{1-\beta}{x^2} < 0,$$

$L(x)$ is decreasing in x, and $L(D_1^{SP}) > L\left(\frac{1}{2}\right)$. Therefore

$$H(D_1^{SP},0) - H\left(\frac{1}{2},0\right) - [c(1)-c(2)] = L(D_1^{SP}) - L\left(\frac{1}{2}\right) > 0.$$

■

B Numerical Methods for Finding Equilibria

B.1 Homotopies.

This Appendix provides details of our implementation of the homotopy algorithm using the example of how we use a sequence of homotopies to try to enumerate the number of equilibria that exist for different values of (ρ, δ) for given values of τ. Our implementation of other homotopies, for example, by varying τ, is similar to a single step in this sequence. We describe the procedure assuming that $M = 30$ and that we are using the price and value formulation of the equations. Identical procedures apply using the choice probability formulation, except that it is necessary to use the choice probabilities to calculate prices and values in order to check whether the identified solutions are different enough to be labeled as distinct equilibria.

B.1.1 Preliminaries

We identify equilibria at particular gridpoints in (ρ, δ) space. We specify a 201-point evenly-spaced grid for the forgetting rate $\delta \in [0, 0.2]$ and a 41-point evenly-spaced grid for the learning progress ratio $\rho \in [0.6, 1]$. The state space of the game is defined by a (30×30) grid of values of the know-how of each firm.

B.1.2 System of Equations Defining Equilibrium

An MPNE is defined by a system of equations (one VS^* equation (text equation (17)) for each of 900 states and one p^* equation (text equation (6)) for each of 900 states. The grouping of all of these equations is denoted F.

B.1.3 Homotopy Algorithm: Overview

The idea of the homotopy is to trace out an equilibrium correspondence as one of the parameters of interest is changed, holding the others fixed. Starting from any equilibrium, the numerical algorithm traces a path where a parameter (such as δ), and the vectors $VS^*(\mathbf{e})$ and $p^*(\mathbf{e})$ are changed together so that the equations F continue to hold, by solving a system of differential equations. The differential equation solver does not return equilibria exactly at the gridpoints so it is necessary to interpolate between the solutions returned by the solver. Homotopies can be run starting from different equilibria and varying different parameters. When these different homotopies return interpolated solutions at the same gridpoint it is necessary to define a numerical rule for when two different solutions should be counted as different equilibria.

B.1.4 Procedure Details

Step 1: Finding Equilibria for $\delta = 0$. The first step is to find an equilibrium (i.e., a solution to the 1,800 equations) for $\delta = 0$ for each value of ρ on the grid. There will be a unique MPNE for $\delta = 0$, as, in this case, movements through the state space are unidirectional, so that the state will eventually end up in the state (M, M) where no more learning is possible.

We solve for an equilibrium using the Levenberg-Marquardt algorithm implemented using `fsolve` in MATLAB, where we supply analytic gradients for each equation. The solution for the previous value of ρ are used as starting values. To ensure that the solutions are precise, we use a tolerance of 10^{-7} for the sum of squared values of each equation, and a relative tolerance of 10^{-14} for the price and value variables that we are solving for.

Step 2: δ-Homotopies. Using the notation of Besanko, Doraszelski, Kryukov, and Satterthwaite (2010), we explore the correspondence

$$F^{-1}(\rho) = \{(\mathbf{V}^*, \mathbf{p}^*, \delta) | F(\mathbf{V}^*, \mathbf{p}^*; \rho, \delta) = \mathbf{0}, \quad \delta \in [0, 1]\},$$

The homotopy approach follows the correspondence as a parameter, s, changes (in our analysis, s could be δ, ρ or τ). Denoting $\mathbf{x} = (\mathbf{V}^*, \mathbf{p}^*)$, $F(\mathbf{x}(s), \delta(s), \rho) = \mathbf{0}$

can be implicitly differentiated to find how $\mathbf{x}$ and δ must change for the equations to continue to hold as s changes.

$$\frac{\partial F(\mathbf{x}(s), \delta(s), \rho)}{\partial \mathbf{x}} \mathbf{x}'(s) + \frac{\partial F(\mathbf{x}(s), \delta(s), \rho)}{\partial \delta} \delta'(s) = \mathbf{0}$$

where $\frac{\partial F(\mathbf{x}(s), \delta(s), \rho)}{\partial \mathbf{x}}$ is a (1,800 x 1,800) matrix, $\mathbf{x}'(s)$ and $\frac{\partial F(\mathbf{x}(s), \delta(s), \rho)}{\partial \delta}$ are both (1,800 x 1) vectors and $\delta'(s)$ is a scalar. The solution to these differential equations will have the following form, where $y_i'(s)$ is the derivative of the ith element of $\mathbf{y}(s) = (\mathbf{x}(s), \delta(s))$,

$$y_i'(s) = (-1)^{i+1} \det \left(\left(\frac{\partial F(\mathbf{y}(s), \rho)}{\partial \mathbf{y}} \right)_{-i} \right)$$

where $_{-i}$ means that the ith column is removed from the (1,801 x 1,801) matrix $\frac{\partial F(\mathbf{y}(s), \rho)}{\partial \mathbf{y}}$.

To implement the path-following procedure, we use the FORTRAN routine FIXPNS from HOMPACK90, with the ADIFOR 2.0D automatic differentiation package used to evaluate the sparse Jacobian $\frac{\partial F(\mathbf{y}(s), \rho)}{\partial \mathbf{y}}$ and the STEPNS routine is used to find the next point on the path.[20,21]

The FIXPNS routine will return solutions at values of δ that are not equal to the gridpoints. Therefore we adjust the code so that after *each* step, the algorithm checks whether a gridpoint has been passed and, if so, the routine ROOTNX is used to calculate the equilibrium at the gridpoint, using information on the solutions at either side.[22]

The time taken to run a homotopy is usually between one hour and seven hours,

[20]STEPNS is a predictor-corrector algorithm where hermetic cubic interpolation is used to guess the next point, and an iterative procedure is then used to return to the path.

[21]For details of the HOMPACK subroutines, please consult manual of the algorithm at `https://users.wpi.edu/~walker/Papers/hompack90,ACM-TOMS_23,1997,514-549.pdf`.

[22]It can happen that the ROOTNX routine stops prematurely so that the returned solution is not exactly at the gridpoint value of δ. We do not use the small proportion of solutions where the difference is more than 10^{-6}. Varying this threshold does not affect the reported results. We also need to decide whether the equations have been solved accurately enough so that the values and strategies can be treated as equilibria. The criteria that we use is that solutions where the value of each equation residual should be less than 10^{-10}. Otherwise, the solution is rejected. In practice, the rejected solutions typically have residuals that are much larger than 10^{-10}.

when it is run on UMD's BSWIFT cluster (a moderately sized cluster for the School of Behavioral and Social Sciences).

Step 3: Enumerating Equilibria. Once we have collected the solutions at each of the (ρ, δ) gridpoints we need to identify which solutions represent distinct equilibria, taking into account that small differences may arise because of numerical differences that are within our tolerances. For this paper, we use the rule that solutions count as different equilibria if at least one element of the price vector differs by more than 0.001.

Step 4: ρ-Homotopies. With a set of equilibria from the δ-homotopies in hand, we can perform the next round of our criss-crossing procedure which alternates ρ-homotopies and δ-homotopies, which we run in both directions (e.g., decreasing ρ as well as increasing ρ). We use equilibria found in the last round as starting points.[23]

This second round of homotopies can also help us to deal with gridpoints where the first round δ-homotopies identify no equilibria because a homotopy run stops (or takes a long sequence of infinitesimally small steps). As noted by Besanko, Doraszelski, Kryukov, and Satterthwaite (2010) (p. 467), the homotopies may stop if they reach a point where the evaluated Jacobian $\frac{\partial F(\mathbf{y}(s),\rho)}{\partial \mathbf{y}}$ has less than full rank. Suppose, for example, that the δ-homotopy for $\rho = 0.8$ stops at $\delta = 0.1$, so we have no equilibria for δ values above 0.1. Homotopies that are run from gridpoints where we did find equilibria with higher values of δ and higher or lower values of ρ may fill in some of the missing equilibria.

Step 5: Repeat Steps 3, 2 and 4 to Identify Additional Equilibria Using New Equilibria as Starting Points. We use the procedures described in Step 3

[23] In practice, using all new equilibria could be computationally prohibitive. We therefore use an algorithm that continues to add new groups of 10,000 starting points when we find that using additional starting points yields a significant number of equilibria that have not been identified before. We have experimented with different rules, and have found that alternative algorithms do not find noticeably more equilibria, across the parameter space, than the algorithm that we use.

to identify new equilibria at the gridpoints. These new equilibria are used to start new sets of δ-homotopies, which in turn can identify equilibria that can be used for new sets of ρ-homotopies. This iterative process is continued until the number of additional equilibria that are identified in a round has no noticeable effect on the heatmaps which show the number of equilibria. For the Besanko, Doraszelski, Kryukov, and Satterthwaite (2010), $\tau = 0$ case, this happens after 8 rounds.

B.2 Method for Finding Equilibria Based on Three Reformulated Equations in the $M = 3$ Model.

We now describe the alternative method that we use to identify equilibria when $M = 3$.

As described in the text, the equilibrium conditions can be reformulated in terms of the probability that seller 1 is chosen in each state. If we restrict ourselves to symmetric equilibria then, together with the restriction that $D_1(e_1, e_2) = 1 - D_1(e_2, e_1)$, then there are just three unknown probabilities. We will use $D_1(1, 2)$, $D_1(1, 3)$ and $D_1(2, 3)$. The equilibrium equations for these three states are:

$$\sigma \log \left(\frac{1}{D_1^*(e_1, e_2)} - 1 \right) - p_1^*(e_1, e_2) + p_2^*(e_1, e_2) = 0, \tag{22}$$

and, from text Section 3.2,

$$\mathbf{p_1} = \mathbf{\Phi}(\mathbf{D_1}) + \mathbf{c_1} - \beta(\mathbf{Q_1} - \mathbf{Q_2})(\mathbf{I} - \beta\mathbf{Q_2})^{-1}[\mathbf{D_1} \circ \mathbf{\Phi}(\mathbf{D_1})]. \tag{23}$$

in vector form, so that we can substitute prices to express the equations (22) in terms of choice probabilities only.

We proceed in the following steps for a given (ρ, δ, τ) combination.

Step 1. Define a grid of possible values for $D_1(1, 2)$ and $D_1(1, 3)$. For each, we use a vector [1e-10, 1e-9, 1e-7, 1e-6, 1e-5, (0.0001:(0.9999-0.0001)/200:0.9999), 1-1e-5, 1-1e-6, 1-1e-7, 1-1e-8, 1-1e-9, 1-1e-10].

Step 2. For every combination on the grid, find for the value of $D_1(2,3)$ which solves the equilibrium equation for state (2,3), and record the values of the equations (22) for states (1,2) and (1,3), in matrices $M(1,2)$ and $M(1,3)$.[24]

Step 3. Use MATLAB `contour` command to define the shapes where the $M(1,2)$ and $M(1,3)$ surfaces are equal to zero.

Step 4. Count all of the intersections of these curves, using the user-defined MATLAB function `InterX` command.[25]

Of course, the contours are calculated using interpolation so the solutions are therefore not quite exact. Therefore,

Step 5. Using the solutions from the contour intersections as starting points, solve the equilibrium equations using `fsolve`.

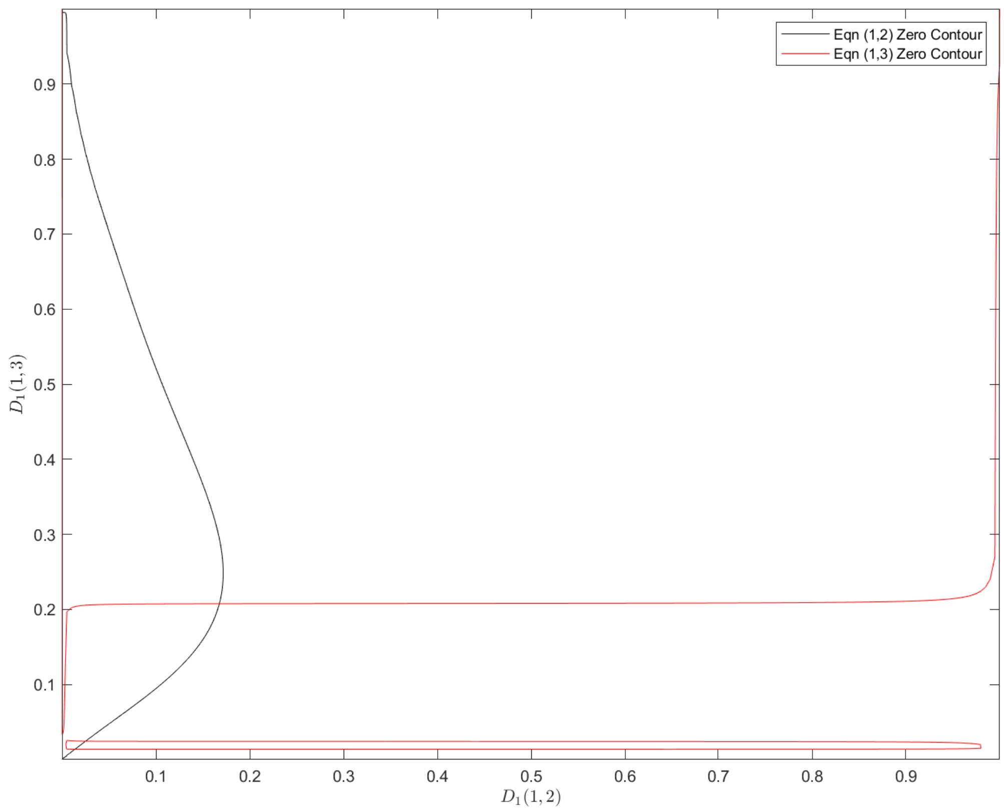
Step 6. Count the number of solutions where at least one choice probability is different from all of the other equilibria by at least 5e-4.

To give a sense of the procedure, consider the parameters $\rho = 0.1$, $\delta = 0.05$ and $\tau = 0.0$. Figure B.2 shows the contour plot. The three intersections between the black and red lines in the bottom left of the figure identify equilibria.

[24] We have not been able to prove uniqueness, but all of the the examples we have looked at there is a unique solution.

[25] `https://www.mathworks.com/matlabcentral/fileexchange/22441-curve-intersections`.

Figure B.2: Illustration of the Contour Plot for $\rho = 0.1$, $\delta = 0.05$ and $\tau = 0$.



C Additional Analysis for the $M = 3$ Model

This Appendix provides some additional analysis for the $M = 3$ model, including for results that are mentioned briefly in the text.

C.1 Equilibrium Strategies and Outcomes for Polar Cases for $\rho = 0.3$ and $\delta = 0.03$.

We use $\rho = 0.3$ and $\delta = 0.03$ as our example parameters in the $M = 3$ model. Table C.2 shows equilibrium prices, sale probabilities and welfare outcomes for (i) the social planner solution, (ii) the equilibrium when $\tau = 0$, and (iii) the equilibrium when $\tau = 1$. The table also shows the probabilities that the industry is in each state after 4 periods (state (3,3) cannot be reached) and 32 periods.

C.2 TS^{PDV}, HHI^{32} and Dynamic Incentives for Alternative ρ and δ.

Our $M = 3$ analysis in the text uses $\rho = 0.3$ and $\delta = 0.03$ as an illustration. We observe that TS^{PDV}, HHI^{32} and the value of dynamic incentives increase as τ increases from zero, with TS^{PDV} and the value of dynamic incentives displaying inverted-U relationships with τ, and concentration a monotonically increasing relationship. The non-monotonic path of dynamic incentives reflects how leads tend to be short-lived when $\tau = 0$, but last longer, so that the value of attaining a lead can increase, as τ rises even though the seller's share of surplus is falling. The non-monotonic path of TS^{PDV} reflects how the social planner would choose more concentration than produced by the $\tau = 0$ equilibrium but less than the equilibrium with $\tau = 1$.

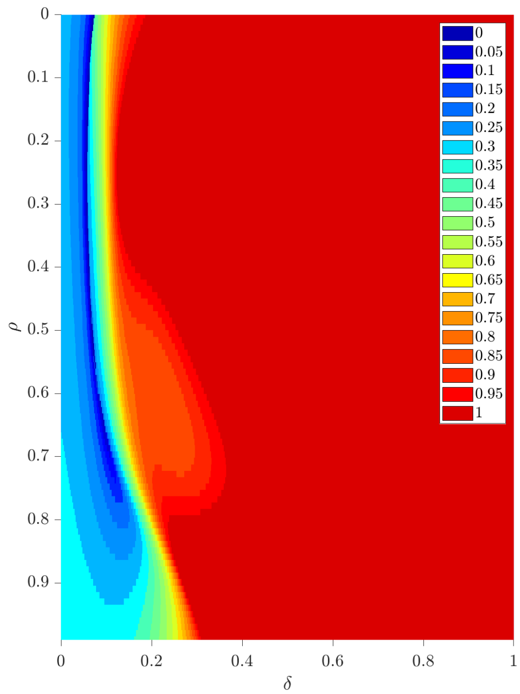
Figure C.2(a)-(c) show the values of τ that maximize TS^{PDV}, HHI^{32} and the PDVs of seller dynamic incentives for all possible ρ and δ combinations using 0.05 steps of τ from 0 to 1. We use the maximum value of the outcome for the small set of (ρ, δ, τ) parameters with multiple equilibria. This choice does affect the value of τ that maximizes the statistic for some (ρ, δ) combinations (see Appendix C.3 for an

Table C.2: $M = 3$ with $\rho = 0.3$ and $\delta = 0.03$: Strategies and Outcomes in Polar Cases. In these tables, firm 2 is assumed to be the leader.

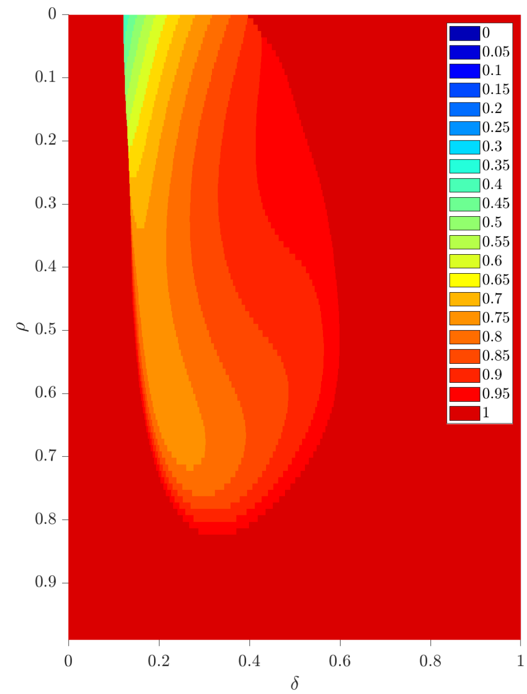
	(a) Social Planner			**(b) $\tau = 0$ Equilibrium**			**(c) $\tau = 1$ Equilibrium**		
	Laggard State Firm 1			Laggard State Firm 1			Laggard State Firm 1		
e_1	1	2	3	1	2	3	1	2	3
c_1	10	3	1.483	10	3	1.483	10	3	1.483
Δ	0.03	0.059	0.087	0.03	0.059	0.087	0.03	0.059	0.087
Prices and Laggard Sale Probability									
$e_2 = 1$	$p_1 = 10$ $p_2 = 10$ $D_1 = 0.5$			$p_1 = 3.378$ $p_2 = 3.378$ $D_1 = 0.5$			$p_1 = 10$ $p_2 = 10$ $D_1 = 0.5$		
$e_2 = 2$	$p_1 = 10$ $p_2 = 3$ $D_1 = 0.084$	$p_1 = 3$ $p_2 = 3$ $D_1 = 0.5$		$p_1 = 4.355$ $p_2 = 3.274$ $D_1 = 0.253$	$p_1 = 2.918$ $p_2 = 2.918$ $D_1 = 0.500$		$p_1 = 10$ $p_2 = 3$ $D_1 = 0.001$	$p_1 = 3$ $p_2 = 3$ $D_1 = 0.500$	
$e_2 = 3$	$p_1 = 10$ $p_2 = 1.483$ $D_1 = 0.114$	$p_1 = 3$ $p_2 = 1.483$ $D_1 = 0.58$	$p_1 = 1.483$ $p_2 = 1.483$ $D_1 = 0.5$	$p_1 = 4.344$ $p_2 = 3.395$ $D_1 = 0.279$	$p_1 = 3.059$ $p_2 = 3.225$ $D_1 = 0.542$	$p_1 = 3.339$ $p_2 = 3.339$ $D_1 = 0.5$	$p_1 = 10$ $p_2 = 1.483$ $D_1 = 0.000$	$p_1 = 3$ $p_2 = 1.483$ $D_1 = 0.180$	$p_1 = 1.483$ $p_2 = 1.483$ $D_1 = 0.5$
4 Period Probability Distribution									
$e_2 = 1$	4.57E-05			8.18E-05			2.72E-05		
$e_2 = 2$	0.0070	0.0190		0.0087	0.050		0.0060	0.0001	
$e_2 = 3$	0.8233	0.1506	0	0.5741	0.3670	0	0.9930	0.0009	0
32 Period State Probability Distribution									
$e_2 = 1$	4.30-09			7.32E-09			1.05E-12		
$e_2 = 2$	2.63E-05	0.0012		1.57E-05	0.00086		6.28E-07	1.97E-05	0
$e_2 = 3$	0.0642	0.1458	0.7888	0.0159	0.1556	0.8277	0.9955	0.0019	0.0026
Expected Welfare Outcomes									
	TS	**CS**	**PS**	**TS**	**CS**	**PS**	**TS**	**CS**	**PS**
PDV	-38.342	-	-	-38.742	-56.169	17.427	-40.702	-40.702	0
4 period	-2.051	-	-	-2.622	-2.797	0.174	-1.492	-1.492	0
32 period	-1.006	-	-	-0.959	-2.621	1.661	-1.481	-1.481	0

Figure C.2: Panels (a)-(c): Values of τ Maximizing TS^{PDV}, HHI^{32} and the PDV of Seller Dynamic Incentives in an $M = 3$ Model. For values of (ρ, δ) with multiple equilibria we use the equilibrium that maximizes the value of the statistic. Panel (d) shows equilibrium, for various τ, and social planner HHI^{32} as a function of δ when $\rho = 0.3$.

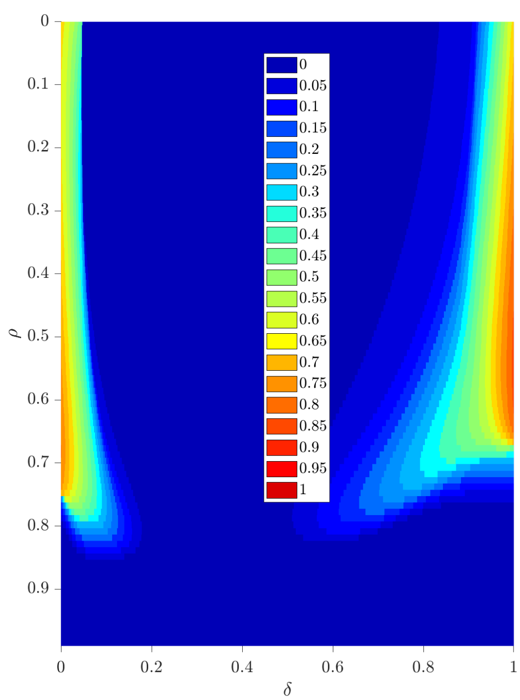
(a) TS^{PDV}



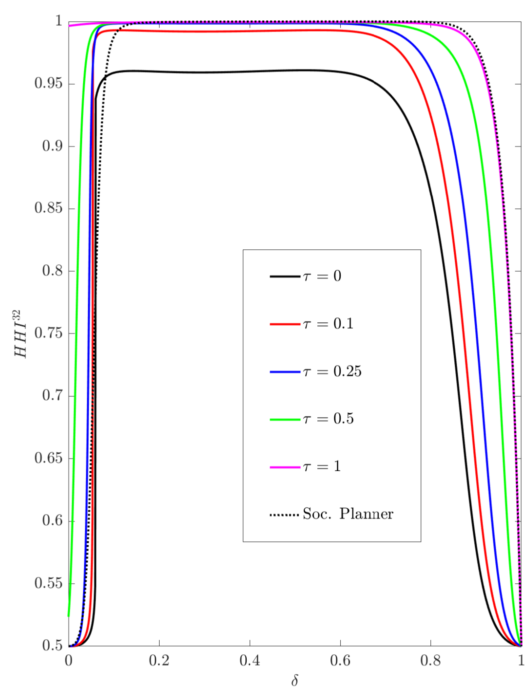
(b) HHI^{32}



(c) PDV of AB+AD Incentives



(d) HHI^{32}



example).

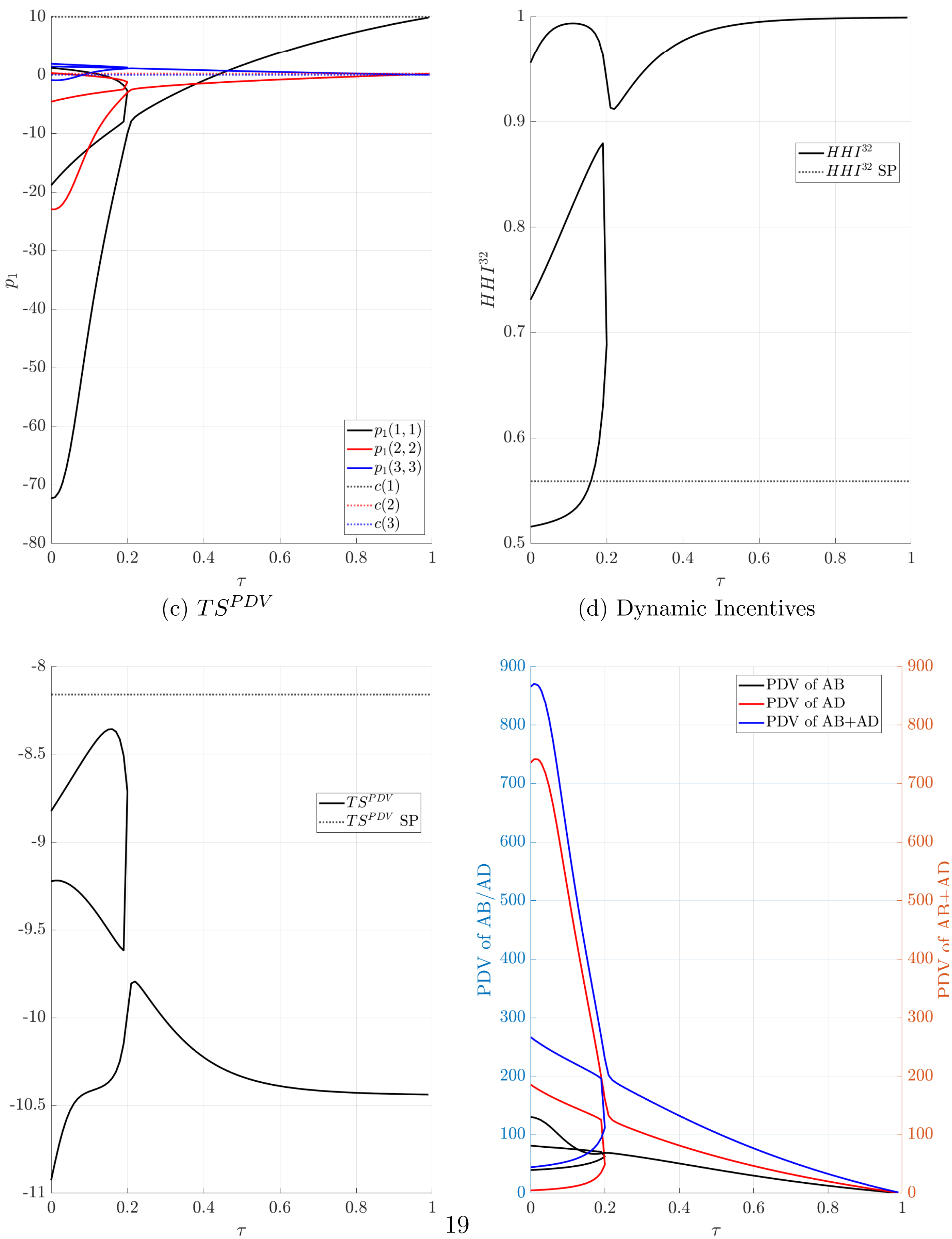
The pattern that raising τ increases concentration is fairly general, and in the cases where HHI is maximized for $\tau < 1$, the level of concentration when $\tau \approx 1$ is exceptionally high. However, the non-monotonicity of TS^{PDV} and dynamic incentives is only a systematic pattern when δ is small and, in the case of dynamic incentives, LBD effects are also at least moderately important.

The patterns change when δ is larger. Panel (d) shows what happens to HHI^{32} in equilibrium, for various τ, and for the social planner, as δ varies and $\rho = 0.3$. For $\delta > 0.1$, the social planner solution involves a single firm making almost every sale, until $\delta > 0.8$, at which point forgetting is so likely that the firm making a sale is unlikely to lower its costs. Equilibria when $0.1 \leq \delta \leq 0.8$ also lead to high concentration, although slightly less than the social optimum due to the leader charging a markup. For $\delta \geq 0.1$, welfare tends to be maximized when $\tau = 1$, consistent with $\tau = 1$ maximizing the probability that a firm with the lowest possible cost will make the sale.

For a wide-range of parameters, dynamic incentives are maximized when $\tau = 0$. This reflects how concentration is high, and leads tend to last a long time, even when $\tau = 0$ for $\delta > 0.1$. When leads last a long time, the lead lengthening effect cannot increase dynamic incentives significantly, and is dominated by how increasing τ shrinks sellers' future profits. When ρ is high, competition is always fairly symmetric, and increasing τ does not cause leads to lengthen very much so that the lead-lengthening effect is also small.[26]

Figure C.3: Prices in Symmetric States, HHI^{32}, TS^{PDV} and Dynamic Incentives along τ-homotopy paths when $\rho = 0.02$ and $\delta = 0.058$.

(a) Prices in Symmetric States (b) HHI^{32} (c) TS^{PDV} (d) Dynamic Incentives



C.3 Analysis for $\rho = 0.02$ and $\delta = 0.058$: Parameters with Multiple Equilibria.

$\rho = 0.3$ and $\delta = 0.03$ support a single equilibrium for all τ. In this subsection, we show how changing τ changes outcomes and dynamic incentives for $\rho = 0.02$ and $\delta = 0.058$. For these parameters, multiplicity exists for $\tau \leq 0.202$. $\rho = 0.02$ implies that production costs fall from 10 to 0.2 when know-how increases from state 1 to 2, i.e., learning-by-doing effects are extreme and sellers in states 2 and 3 have almost the same costs, but a state 2 seller is at risk of experiencing a very large cost increase (and cost disadvantage) if it does not make a sale. As δ is not too large, the social planner would prefer to have both firms in states 2 and 3, with $D_1^{SP}(2,3) \approx 0.8$ so that both firms are very likely to be low cost in the next period.

Figure C.3 shows the values of equilibrium HHI^{32}, TS^{PDV}, prices in symmetric states, and the PDVs of AB and AD incentives along τ-homotopy paths that start from the three equilibria that are identified when $\tau = 0$. The $\tau = 0$ equilibrium that is on the path that continues all of the way to $\tau = 1$ has a pronounced diagonal trench in the sense of having very low, and below cost, "aggressive" prices in all symmetric states, high equilibrium concentration ($D_1^*(1,2) \approx 0.07$), and large AD incentives. In contrast, prices in asymmetric states are higher (for example, in (3,1) prices are 6.1 and 10.5), reflecting how a laggard has limited incentive to try to catch up when symmetric competition is so fierce. This equilibrium minimizes total surplus as the costs associated with reduced variety are larger than the benefits of lower expected costs. The two other $\tau = 0$ equilibria involve firms setting higher prices in symmetric states, including above cost prices in state (3,3). In one of the equilibria concentration is lower than the social planner would choose.

As τ increases, the benefits of achieving a lead in the diagonal trench equilibrium decrease, so that there is a range of τ, between 0.1 and 0.22, where equilibrium

[26] For example if $\rho = 0.9$ and $\delta = 0.03$, increasing τ from 0 to 0.2, causes the lead of a firm in (2,1) to last an expected 2.4, rather than 2.3, periods. The PDV of AD incentives does increase slightly over this range, but this is offset by the value of AB incentives falling by slightly more.

concentration falls, and symmetric prices rise. On the other hand, concentration increases on the loop from the other two $\tau = 0$ equilibria. TS^{PDV} is maximized on the loop path that does not extend past $\tau = 0.202$.

For the $M = 30$ model with illustrative technology parameters ($\rho = 0.75, \delta = 0.023$), we also find three equilibria when $\tau = 0$. Two of them have diagonal trenches, and they are on a loop that does not extend past $\tau = 0.07$. Based on that example, we have been asked by discussants and seminar participants whether diagonal trench equilibria are always on paths that do not continue once τ is large enough, as this would suggest one might be able to view variation in buyer bargaining power as some type of equilibrium selection device. This $M = 3$ example provides a counter-example to this conjecture, and, while diagonal trench equilibria are often eliminated, we have identified similar counterexamples for $M = 30$ as well.

C.4 Introducing an Outside Good and Varying σ.

We follow Besanko, Doraszelski, Kryukov, and Satterthwaite (2010) in assuming that there is no outside good and that σ, which controls the degree of product differentiation, equals 1. In this appendix we examine how far our conclusions about the effects of bargaining power on outcomes depend on these assumptions, assuming $\rho = 0.3$ and $\delta = 0.03$.

C.4.1 Outside Good.

When the buyer is able to choose not to purchase, sellers face additional competition which will constrain markups. One can construct intuitions where this either weakens the incentive of a firm to establish a lead over its rival (as it will reduce the return to establishing a lead) or strengthens it (as competition from the outside good may make it even harder for a laggard to catch up).

We introduce an outside good by assuming that, in every period, buyers have a third option, with indirect utility $v - p_0 + \varepsilon_0$, where p_0 is an exogenous parameter that we can use to control the attractiveness of the outside good. Besanko, Doraszelski, and Kryukov (2014), who consider a model where only a single seller may be active, allow an outside good with a baseline $p_0 = 10$.

Figure C.4(a) shows the equilibrium values of HHI^{32} for $\tau = 0$, 0.2 and 0.5 as a function of p_0. Figure C.4(c) and (e) show the PDV of AB and AD incentives. The BDKS model can be viewed as the limiting case where $p_0 \to \infty$ (i.e., we extend the right-hand edge of the figure further to the right). The effects of the outside good are small unless $p_0 < 7$, in which case dynamic incentives increase and concentration (reflecting the dominance of one seller over its rival) increases slightly. However, for $p_0 \approx 10$, all measures are very similar to those when an outside good is assumed not to exist, and introducing an outside good has little effect on the comparisons between different τs.

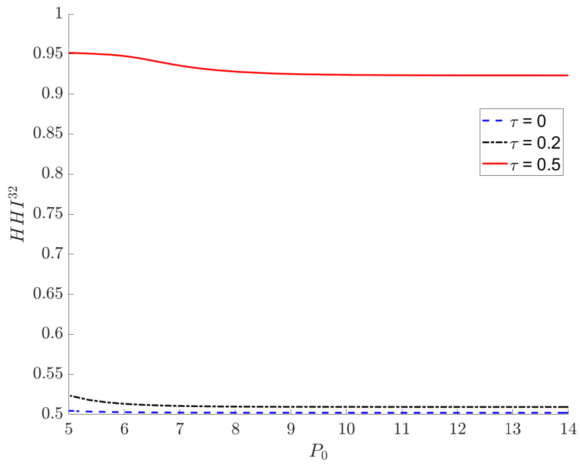
C.4.2 Changing Product Differentiation.

As σ increases, it becomes more likely that, in any period, the buyer will have a strong preference for one of the sellers, so that seller competition is softened. This will tend to make it more likely that purchases will be evenly split across sellers, although the expectation of future sales and softened competition could increase a firm's incentive to lower its own costs.

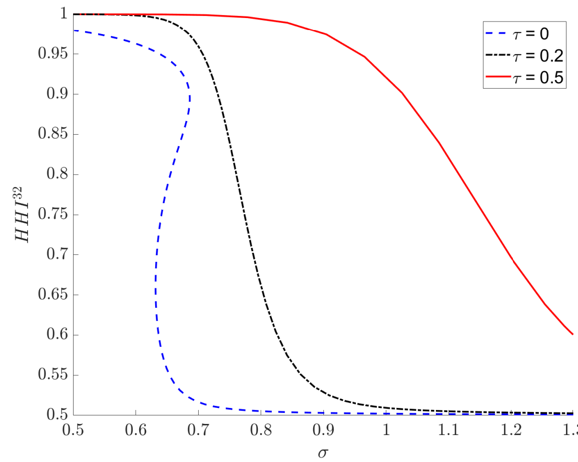
Figure C.4 shows the same statistics as σ is varied, assuming that there is no outside good. Increasing σ from 1 lowers concentration, but does not dramatically change how giving buyers bargaining power affects outcomes. On the other hand, reducing σ to 0.8, or lower, causes equilibrium concentration to increase sharply and can introduce multiple equilibria when $\tau = 0$, illustrated by the $\tau = 0$ path bending back on itself. Higher concentration is associated with leads lasting longer when $\tau = 0$, which tends to lead to dynamic incentives monotonically declining in τ.

Figure C.4: HHI^{32} and the PDV of Dynamic Incentives as a Function of the Exogenous Price of an Outside Good (panels a, c, and e) and Product Differentiation (b, d and f) when $M = 3$, $\rho = 0.3$ and $\delta = 0.03$.

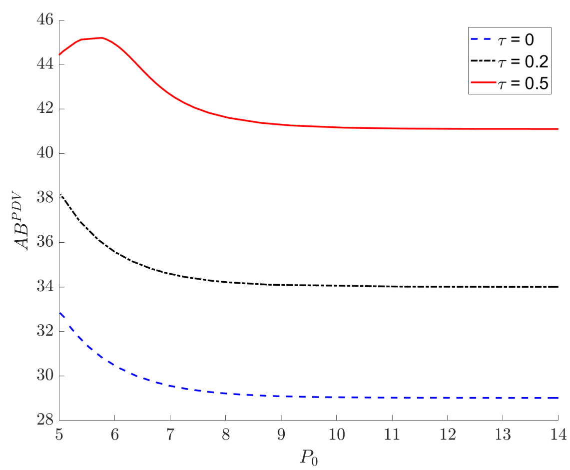
(a) Outside Good: HHI^{32}



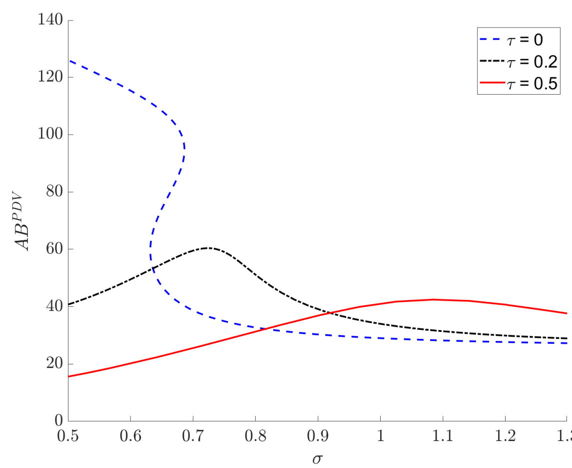
(b) Differentiation: HHI^{32}



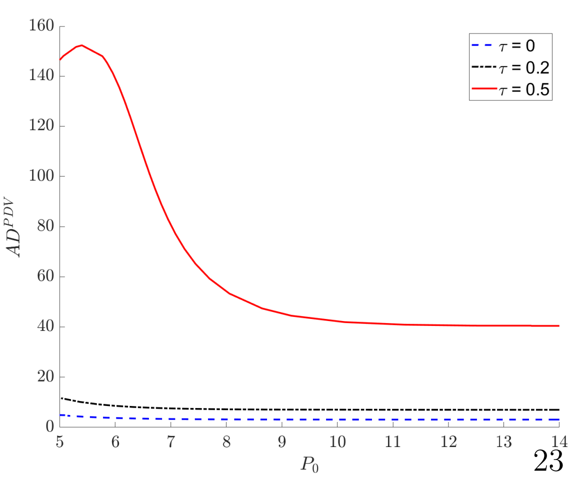
(c) Outside Good: PDV of AB Incentives



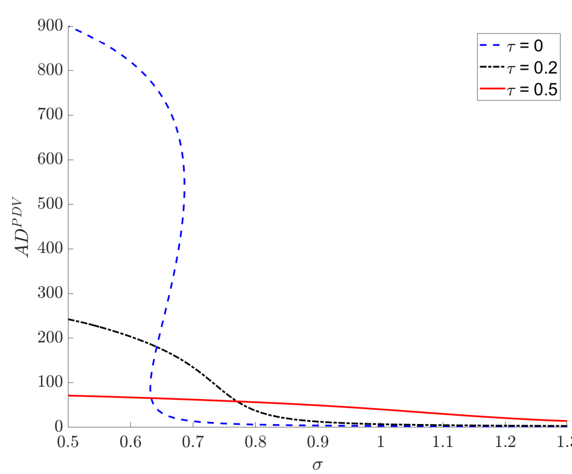
(d) Differentiation: PDV of AB Incentives



(e) Outside Good: PDV of AD Incentives



(f) Differentiation: PDV of AD Incentives



C.5 Subsidies.

We calculate subsidies that could implement the social planner outcome. The text shows that the subsidies that would be optimal if $\rho = 0.3$, $\delta = 0.03$ and $\tau = 0$ lower welfare, relative to the no subsidy equilibrium, if $\tau \geq 0.06$.

We try to investigate whether optimal $\tau = 0$ subsidies lower welfare for other technologies when $\tau > 0$. We specifically consider $\tau = 0.2$ so that sellers have most, but not all, of the bargaining power and an economist might assume that a $\tau = 0$ analysis would provide a reasonable approximation if they do not appreciate how quickly strategies and outcomes change when τ increases from zero.

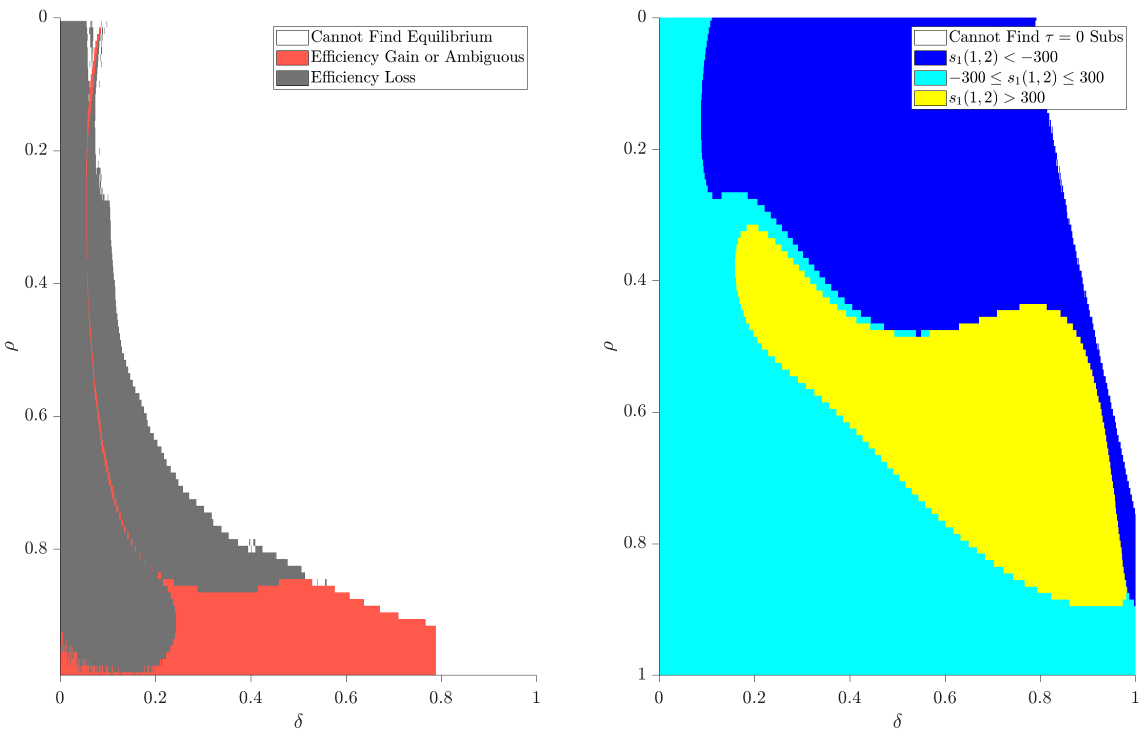
Figure C.5(a) shows that $\tau = 0$ subsidies lower welfare if $\tau = 0.2$ for a wide range of parameters as long as LBD effects are not too limited ($\rho < 0.85$). The red areas indicate cases where there is either an efficiency gain or the existence of multiple equilibria when $\tau = 0.2$ means that we cannot sign whether welfare is increased or reduced. The white area indicates parameters where we cannot solve for $\tau = 0$ subsidies or we cannot solve for the equilibrium when $\tau = 0.2$ when these subsidies are in place.

The fact that the white areas cover so many parameters may seem surprising. It reflects the extreme values of some of the required subsidies or the extreme values of the equilibrium choice probabilities when these subsidies are in place. As illustration, panel (b) shows the level of the subsidy the planner would want to give to a laggard making a sale in state (1,2). For $\delta > 0.1$, the subsidies or taxes can be extremely large. It is also interesting how a small change in ρ can switch the optimal scheme from providing a laggard with a very large subsidy to requiring the laggard to pay a very large tax.

Figure C.5: Optimal $\tau = 0$ Subsidies and Welfare in the $M = 3$ Model. Panel (a) shows, for the values of ρ and δ for which we can identify equilibria, parameters where $\tau = 0$ optimal subsidies increase or decrease welfare when $\tau = 0.2$ (compared to no subsidy equilibria). Panel (b) shows the value of $\tau = 0$ subsidies to the laggard in state (1,2). A negative subsidy is a tax on a laggard sale.

(a) Welfare Effect of $\tau = 0$ Subsidies if $\tau = 0.2$

(b) Level of $\tau = 0$ Laggard Subsidy in State (1,2)



D Additional Analysis for the $M = 30$ Model

D.1 Equilibria for $\tau = 0$, $\rho = 0.75$ and $\delta = 0.023$.

Table D.1 lists strategies in a subset of states for the three equilibria that exist for the illustrative technology parameters when $\tau = 0$. All equilibria have negative prices in the initial state (1,1). The B and C equilibria are characterized by a "diagonal trench" with lower prices when firms are symmetric or almost symmetric. For example, prices in (10,10) are 2.43, whereas prices in (29,10), where costs are weakly lower, are 5.39 and 5.15. The trench in equilibrium B does not extend to the highest levels of know-how, and it has slightly lower $HHI^{1,000}$, as it is more likely that the sellers will be symmetric in the long-run. Equilibrium A, with the lowest $HHI^{1,000}$, has prices that vary less with know-how once both firms reached know-how states 3 or 4. However, the leader sets lower prices when the laggard is in know-how states 1 or 2, so that, at the start of the game, it is more likely that one of the sellers will move down its cost curve more quickly, so that HHI^{32} is higher.

D.2 Concentration and Bargaining Power.

Text section 5 shows, for the illustrative technology parameters, that:

- concentration is similar in dynamic equilibria and in an equilibrium where firms price statically, and concentration is lower than the social planner would choose, when $\tau = 0$.
- static and dynamic equilibria are identical when $\tau = 1$, and concentration is above the level that the social planner would choose.
- with static pricing, concentration increases fairly steadily as τ increases from 0 to 1, whereas, with dynamic behavior, concentration increases sharply (and actually overshoots its $\tau = 1$ level) as τ increases from zero.

Table D.1: Equilibria in the Besanko, Doraszelski, Kryukov, and Satterthwaite (2010) Model for Illustrative Parameters ($\delta = 0.023$, $\rho = 0.75$) and $\tau = 0$ in the $M = 30$ and $m = 15$ Model.

						Eqm. A $HHI^{32} = 0.537$ $HHI^{1,000} = 0.500$		Eqm. B $HHI^{32} = 0.520$ $HHI^{1,000} = 0.516$		Eqm. C $HHI^{32} = 0.520$ $HHI^{1,000} = 0.527$	
e_1	e_2	c_1	c_2	Δ_1	Δ_2	p_1	p_2	p_1	p_2	p_1	p_2
1	1	10.00	10.00	0.0230	0.0230	-0.54	-0.54	-1.63	-1.63	-1.61	-1.61
2	1	7.50	10.00	0.0455	0.0230	4.91	7.21	5.16	7.60	5.15	7.60
2	2	7.50	7.50	0.0455	0.0455	4.22	4.22	0.77	0.77	0.77	0.77
3	1	6.34	10.00	0.0674	0.0230	5.82	8.18	6.56	8.71	6.55	8.70
3	2	6.34	7.50	0.0674	0.0455	4.65	5.46	4.06	5.97	4.05	5.97
3	3	6.34	6.34	0.0674	0.0674	5.11	5.11	1.49	1.49	1.47	1.47
4	1	5.62	10.00	0.0889	0.0230	5.95	8.29	6.67	8.55	6.67	8.54
4	2	5.62	7.50	0.0889	0.0455	4.86	5.85	5.46	7.08	5.46	7.08
4	3	5.62	6.34	0.0889	0.0674	5.08	5.42	3.81	5.53	3.80	5.54
4	4	5.62	5.62	0.0889	0.0889	5.22	5.22	1.72	1.72	1.70	1.70
10	1	3.85	10.00	0.2076	0.0230	5.89	8.16	6.14	7.71	6.13	7.69
10	2	3.85	7.50	0.2076	0.0455	5.05	6.06	5.75	6.43	5.75	6.42
10	3	3.85	6.34	0.2076	0.0674	5.20	5.81	5.80	6.31	5.80	6.31
10	8	3.85	4.22	0.2076	0.1699	5.10	5.20	4.49	5.85	4.49	5.86
10	9	3.85	4.02	0.2076	0.1889	5.11	5.15	3.26	4.56	3.25	4.55
10	10	3.85	3.85	0.2076	0.2076	5.12	5.12	2.47	2.47	2.43	2.43
15	1	3.25	10.00	0.2946	0.0230	5.79	8.05	5.98	7.36	5.97	7.38
15	2	3.25	7.5	0.2946	0.045	5.02	5.93	5.63	6.18	5.62	6.17
15	3	3.25	6.34	0.2946	0.0674	5.22	5.74	5.67	6.02	5.67	6.01
15	10	3.25	3.85	0.2946	0.2076	5.19	5.20	5.40	5.94	5.41	5.95
15	14	3.25	3.34	0.2946	0.2780	5.23	5.21	3.46	4.44	3.43	4.44
15	15	3.25	3.25	0.2946	0.2946	5.24	5.24	3.16	3.16	3.10	3.10
16	16	3.25	3.25	0.3109	0.3109	5.28	5.28	3.24	3.24	3.18	3.18
20	20	3.25	3.25	0.3721	0.3721	5.25	5.25	3.32	3.32	3.20	3.20
22	22	3.25	3.25	0.4007	0.4007	5.25	5.25	3.44	3.44	3.26	3.26
25	25	3.25	3.25	0.4411	0.4411	5.25	5.25	3.90	3.90	3.28	3.28
27	27	3.25	3.25	0.4665	0.4665	5.25	5.25	4.62	4.62	3.34	3.34
28	28	3.25	3.25	0.4787	0.4787	5.25	5.25	4.98	4.98	3.52	3.52
29	1	3.25	10.00	0.4907	0.0230	5.79	8.05	5.63	7.62	5.57	7.46
29	2	3.25	7.50	0.4907	0.0455	5.01	5.91	5.04	5.77	5.07	5.72
29	10	3.25	3.85	0.4907	0.2076	5.23	5.17	5.35	5.17	5.39	5.15
29	15	3.25	3.25	0.4907	0.2946	5.27	5.22	5.45	5.34	5.52	5.33
29	29	3.25	3.25	0.4907	0.4907	5.25	5.25	5.22	5.22	3.98	3.98
30	1	3.25	10.00	0.5024	0.0230	5.79	8.05	5.67	7.66	5.63	7.53
30	2	3.25	7.50	0.5024	0.0455	5.01	5.91	5.10	5.84	5.12	5.81
30	10	3.25	3.85	0.5024	0.2076	5.23	5.17	5.33	5.21	5.35	5.19
30	15	3.25	3.25	0.5024	0.2946	5.27	5.22	5.42	5.35	5.45	5.33
30	29	3.25	3.25	0.5024	0.4907	5.25	5.25	5.30	5.20	4.29	4.60
30	30	3.25	3.25	0.5024	0.5024	5.25	5.25	5.27	5.27	4.77	4.77

Notes: c_i, p_i, Δ_i are the marginal costs, equilibrium price and probability of forgetting for firm i. HHI^∞ is the expected long-run value of the HHI.

To investigate how far these results depend on the specific technology parameters, Figure D.1 shows the level of HHI^{32} under the social planner solution, and the dynamic equilibria when $\tau = 0$, 0.1, 0.25, 0.5 and 1 for the ranges of (ρ, δ) that we consider. When multiple equilibria exist, the maximum HHI^{32} is shown. The existence of multiple equilibria is the reason for the discontinuities in shading in the $\tau = 0$ and $\tau = 0.1$ figures. Figure D.2 shows the level of HHI^{32} under the social planner solution, and the unique equilibria when $\tau = 0$, 0.1, 0.25, 0.5 and 1 with static seller behavior (i.e., sellers' dynamic incentives are set to zero) in each state. Static and dynamic equilibria are identical when $\tau = 1$.

We observe that

- if $\tau = 0$, dynamic and static equilibrium HHI^{32} is generally below the level that the social planner would choose except for high ρ (limited LBD) and δs between 0.025 and 0.1. For this range of δs multiple equilibria are common, and the maximum equilibrium HHI^{32} is low but slightly larger than the social planner would choose.

- if $\tau = 0$, dynamic and static equilibrium HHI^{32}s are generally similar unless $\delta > 0.03$ (which implies $\Delta(m) > 0.37$ and $\Delta(M) > 0.6$ so depreciation rates are quite high). If $\delta > 0.03$, HHI^{32} tends to be larger in dynamic equilibria.

- dynamic equilibrium concentration generally increases sharply as τ increases from zero. Concentration when $\tau = 0.5$ is similar to concentration when $\tau = 1$, although there are examples where HHI^{32} is at a high level but declines slightly as τ increases from 0.5 to 1. Concentration increases more gradually with τ in the static equilibria.

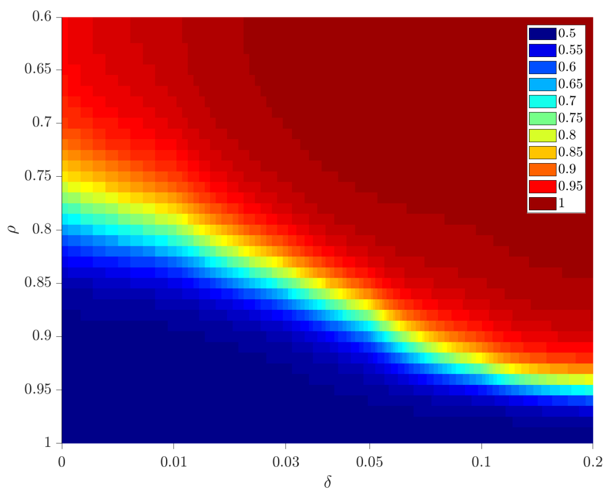
- for $\delta < 0.03$, social planner concentration is somewhere between dynamic equilibrium concentration when $\tau = 0.25$ and dynamic equilibrium concentration when $\tau = 0.5$. This is consistent with our finding that TS^{PDV} is maximized for

τs around 0.3 for these δs (see text Figure 6(a)). For $\delta > 0.03$, concentration when $\tau = 0.5$ is also similar to the socially optimal levels.

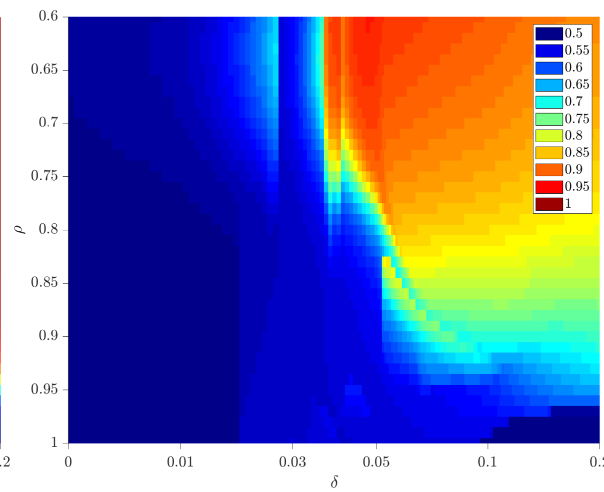
The patterns observed for the illustrative technology parameter results are therefore fairly typical of what we see for other parameters that imply significant LBD effects if know-how depreciation is also limited. The next sub-section provides some analysis for $\rho = 0.95$ and $\delta = 0.03$, which is one example where $\tau = 0$ equilibrium concentration is above the social planner level.

Figure D.1: Expected Value of the HHI^{32} in the Social Planner Solution and Dynamic Equilibria. HHI^{32} value is the maximum across equilibria.

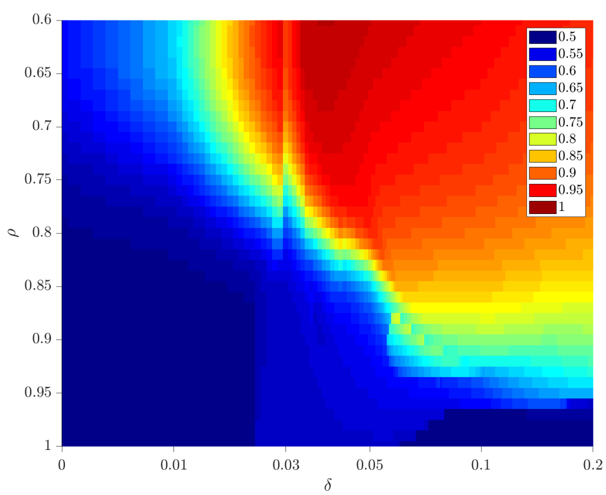
(a) Social Planner



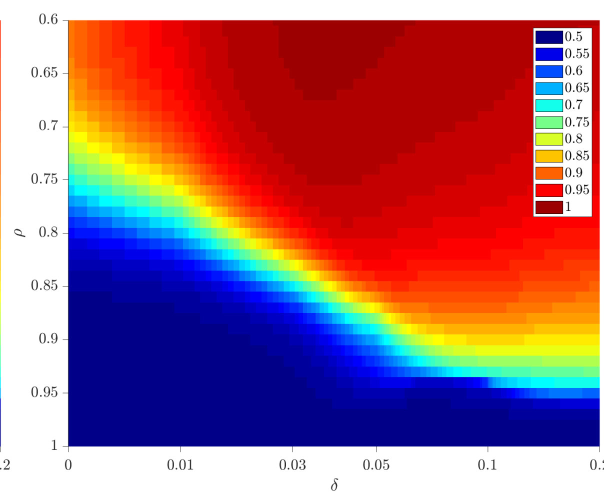
(b) $\tau = 0$



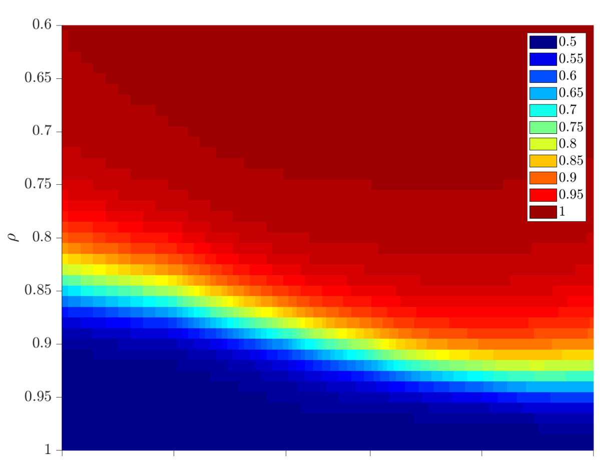
(c) $\tau = 0.1$



(d) $\tau = 0.25$



(e) $\tau = 0.5$



(f) $\tau = 1$

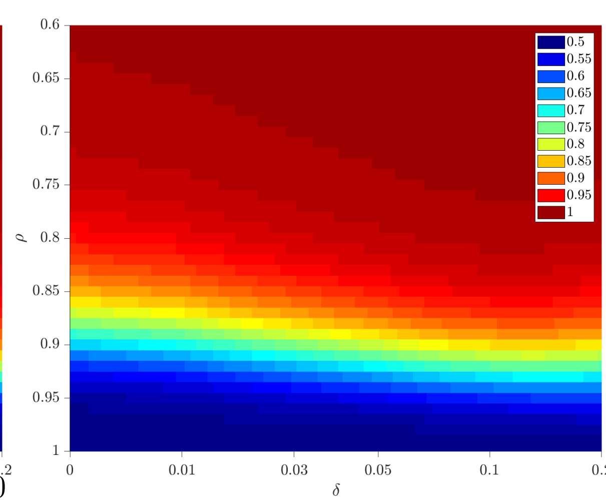
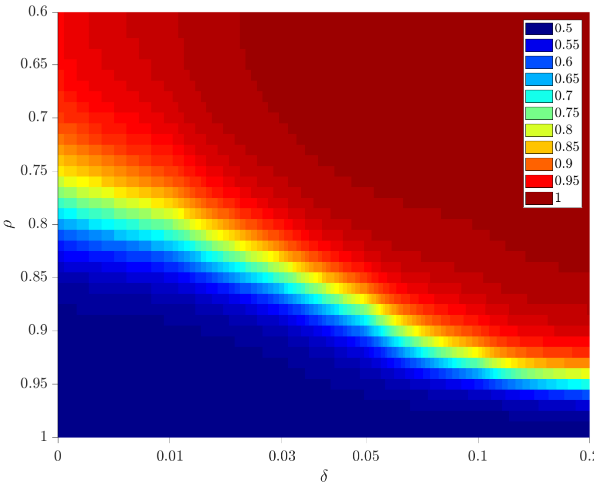
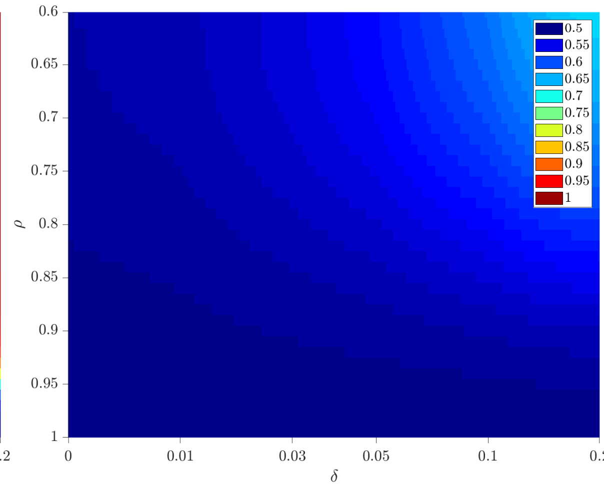


Figure D.2: Expected Value of the HHI[32] in the Social Planner Solution and Equilibria with Static Seller Behavior. All static equilibria are unique.

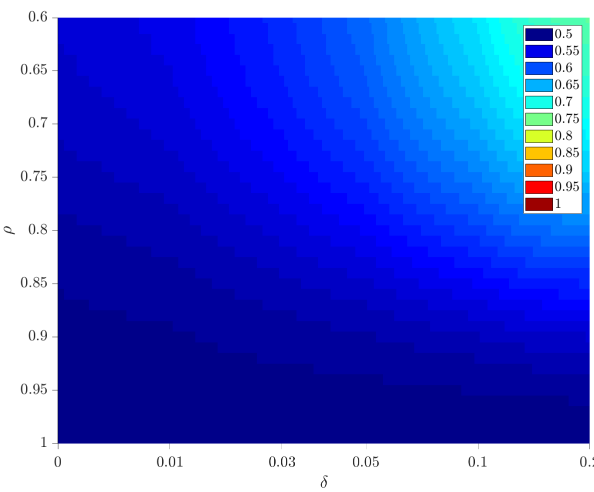
(a) Social Planner



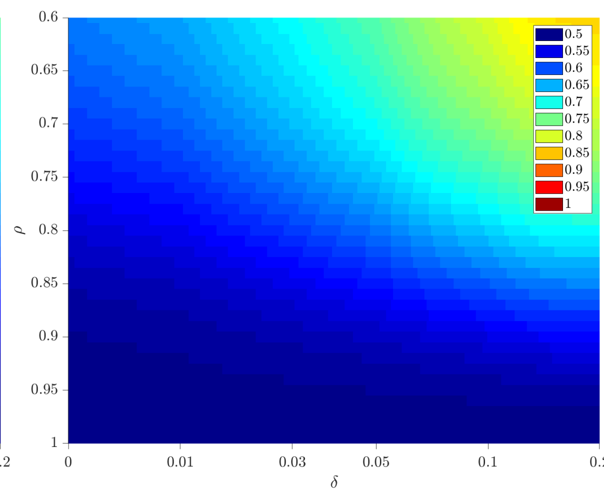
(b) $\tau = 0$ Static



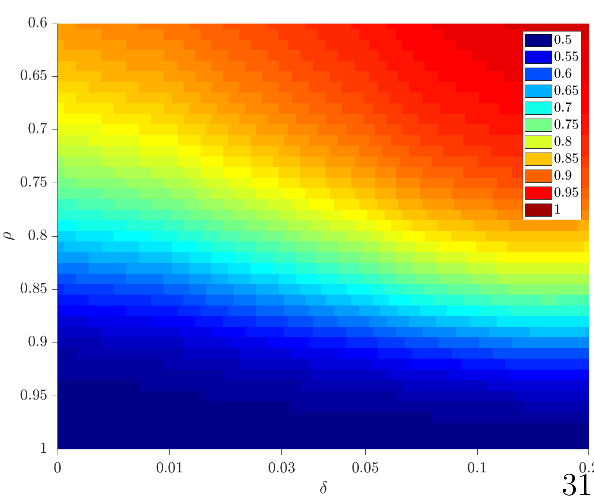
(c) $\tau = 0.1$ Static



(d) $\tau = 0.25$ Static



(e) $\tau = 0.5$ Static



(f) $\tau = 1$ Static

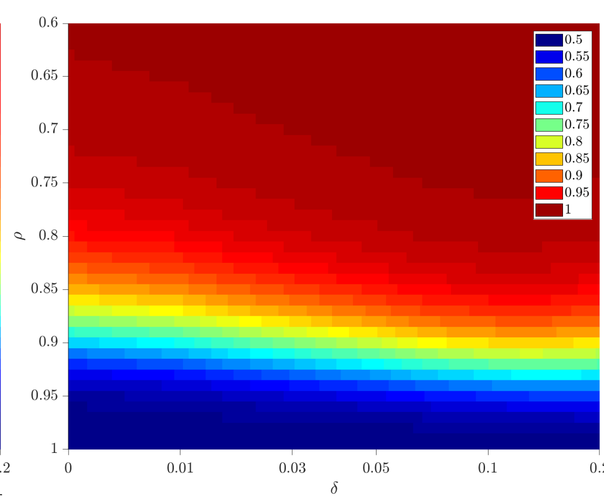
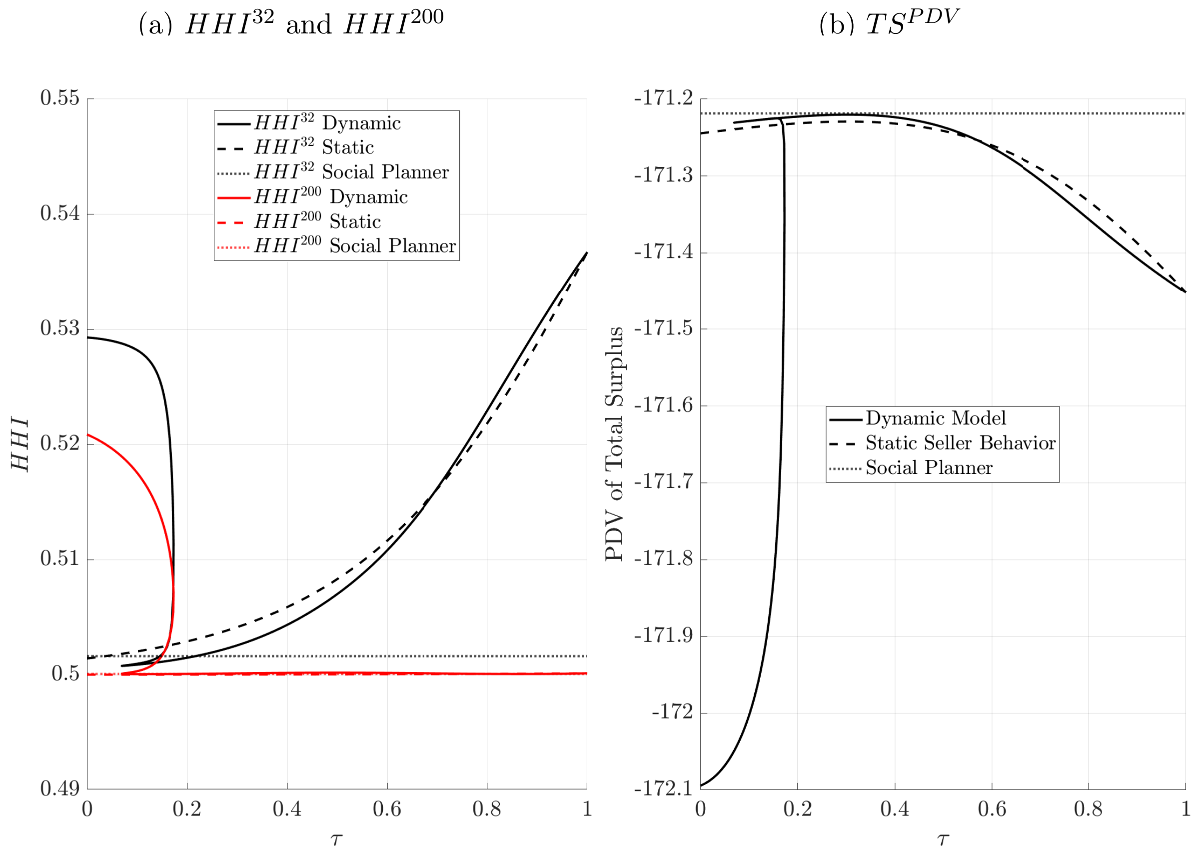


Figure D.3: The Effects of Changing the Allocation of Bargaining Power for $\rho = 0.95$ and $\delta = 0.03$ With No Policies.



D.3 Effects of Bargaining Power For $\rho = 0.95$ and $\delta = 0.03$.

In the text, we use $\rho = 0.75$ and $\delta = 0.023$ as our illustrative technology parameters. For these parameters, market concentration is significantly lower than the social planner would choose when $\tau = 0$. In this subsection, we consider $\rho = 0.95$ and $\delta = 0.03$ as an example of technologies where $\tau = 0$ equilibrium concentration is higher than the social planner would choose. The discounted value of dynamic incentives also declines in τ and multiple equilibria exist for some low τ even though the $\tau = 0$ equilibrium is unique. $\rho = 0.95$ implies limited LBD: know-how can lower production costs by no more than 18%.

Figure D.3 replicates text Figure 7 for our new parameters. The unique equilibrium with $\tau = 0$ has concentration a little above the socially optimal level, and the level that would be generated by static seller pricing. The dynamic equilibrium τ-homotopy path bends back on itself so that there are multiple equilibria for $0.07 \leq \tau \leq 0.175$. In this range, concentration and welfare in the less concentrated equilibria are very close to the socially optimal and static equilibrium levels.[27] For $\tau \geq 0.175$, the equilibrium concentration increases with τ and efficiency declines for $\tau \geq 0.3$.

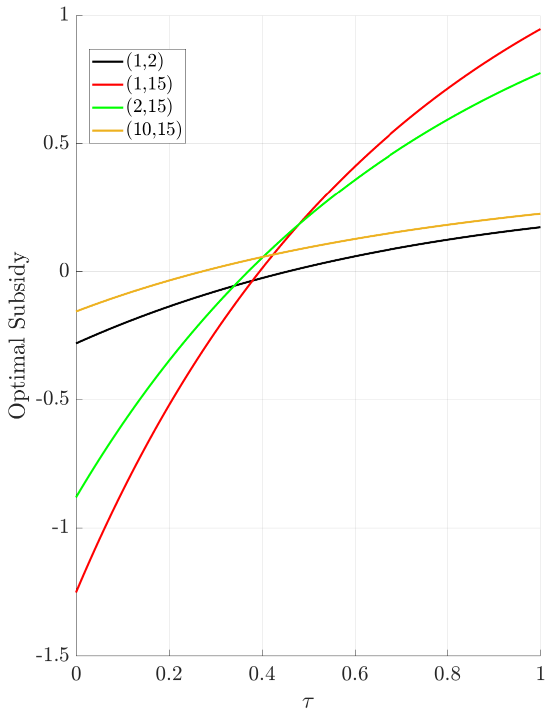
Policies. Figure D.4(a) shows the optimal subsidies that would implement the social planner solution as a function of τ. When $\tau = 0$, the subsidies are much smaller in scale than for the illustrative technology parameters, consistent with equilibrium concentration being closer to the socially optimal level. Even though equilibrium concentration is too high when $\tau = 0$, laggards are taxed when they make a sale. This reflects how, given socially optimal sales probabilities in other states, the laggard would be too likely to make a sale without a tax. Panel (b) shows that the $\tau = 0$ subsidy scheme lowers welfare if $\tau \geq 0.175$. For $0.08 \leq \tau \leq 0.175$ the scheme may increase or decrease welfare depending on which no-subsidy equilibrium is played. Therefore, the conclusion that subsidies that would maximize welfare when $\tau = 0$ can lower welfare for values of τ that are greater than zero but small remain.

Panels (c) and (d) show the effects of our stylized policies to promote competition. We only consider policies that are introduced from the start of the industry. As equilibrium concentration is very low until τ approaches 1, the concentration restriction has almost no effect on welfare or concentration until τ is large. The incentive policies increase welfare when $\tau \approx 0$, and they also increase welfare for $\tau > 0.7$.

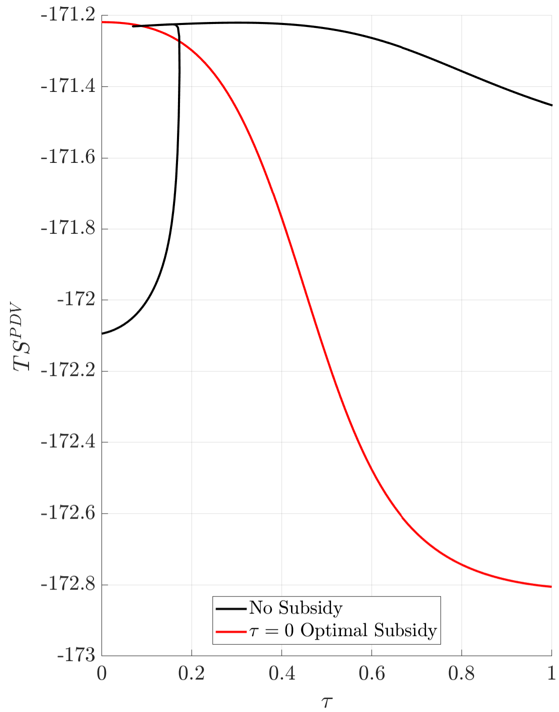
[27] The decline in concentration as τ increases is also associated with the leads of leaders tending to last for fewer periods, which also means that sellers' dynamic AB and AD incentives decline in τ, rather than increasing-then-decreasing as they do for the illustrative parameters.

Figure D.4: Policies and Bargaining Power for $\rho = 0.95$ and $\delta = 0.03$. Subsidies are given to the laggard when it makes a sale. This analysis assumes that policies are introduced at the start of the industry's life.

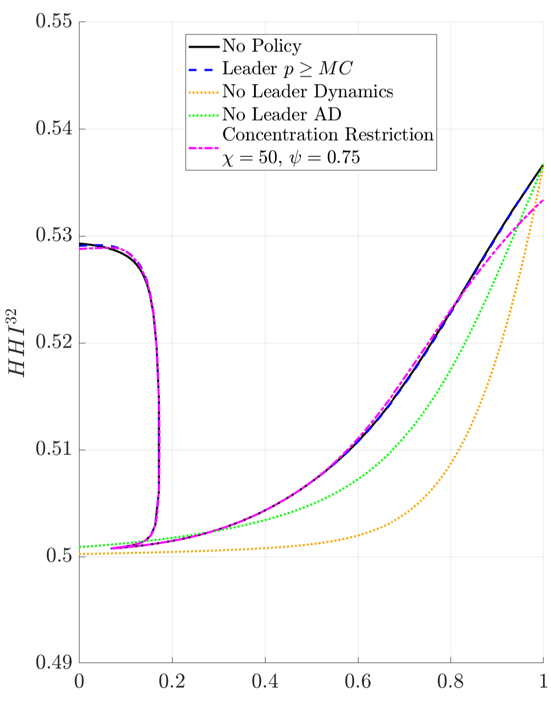
(a) Optimal Subsidies in Selected States



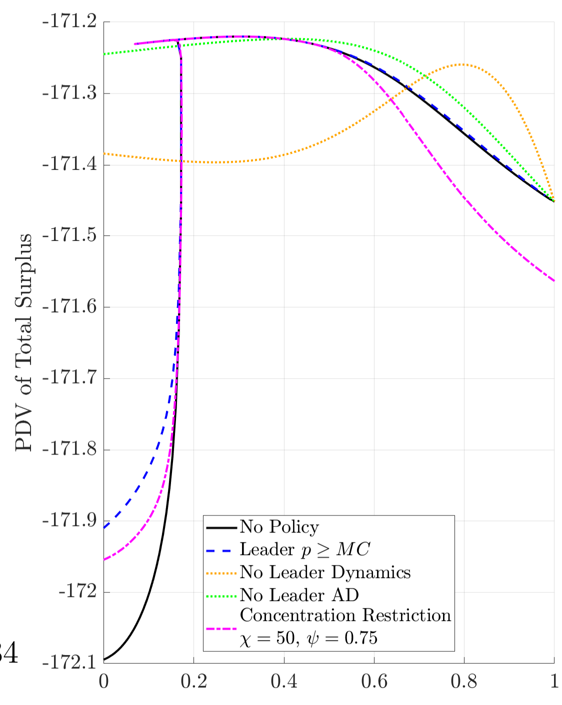
(b) Welfare and $\tau = 0$ Optimal Subsidies



(c) Policies Promoting Competition: HHI^{32}



(d) Policies Promoting Competition: TS^{PDV}



D.4 Alternative Concentration Restriction Policies.

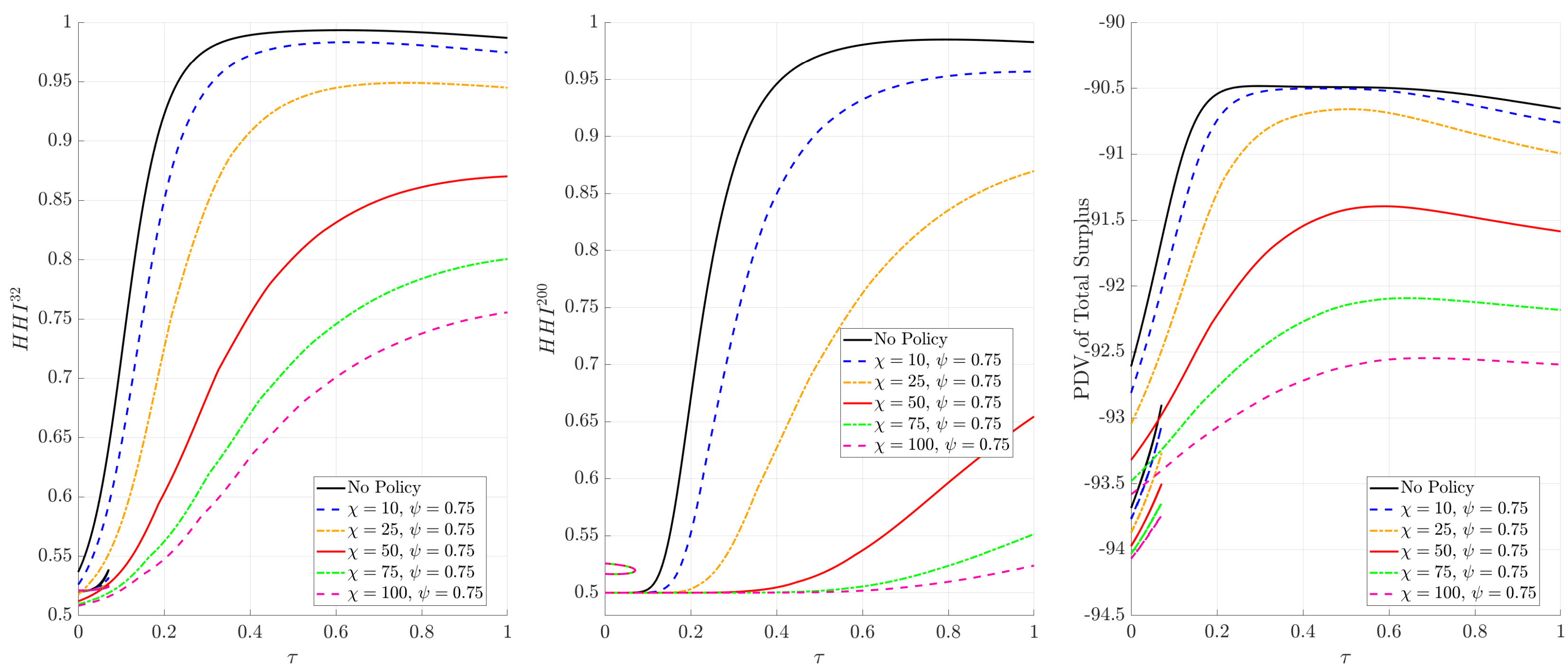
Text section 5.3 shows the effects of a concentration restriction policy, where the leader i has to pay a compliance penalty of $\chi \times \max\{0, D_i - \psi\}$ where $\chi = 50$ and $\psi = 0.75$. As $D_i = 0.5$ when firms are symmetric, and the maximum D_i approaches 1, $\psi = 0.75$ is a natural value to consider. We choose $\chi = 50$ as an example of a policy which lowers concentration but which still provides some probability that a firm will establish a know-how advantage that will lead to the compliance cost being incurred.

Figure D.5 shows how policies with different χs affect concentration and discounted total welfare (recall that we do not count the compliance cost as a total welfare loss) for the illustrative technology parameters, as a function of τ. Consistent with what one would expect, increases in χ lower concentration for values of τ where the share threshold would likely be breached with no policy in effect. For the values of χ that we consider, welfare falls as χ increases, although further analysis identifies that for $\tau \approx 0.4$ the policy can slightly increase TS^{PDV} when χ is slightly greater than zero.[28]

[28]For example, when $\tau = 0.4$, $TS^{PDV} = -90.4880$ when $\chi = 2$, compared to -90.4883 when $\chi = 0$ (no policy).

Figure D.5: Concentration and Welfare for Alternative Compliance Cost Parameters and the Concentration Restriction Policy, Assuming the Illustrative Technology Parameters.

(a) HHI^{32} (b) HHI^{200} (c) TS^{PDV}



References

Besanko, D., U. Doraszelski, and Y. Kryukov (2014): "The Economics of Predation: What Drives Pricing When There is Learning-by-Doing?," *American Economic Review*, 104(3), 868–97.

Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite (2010): "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics," *Econometrica*, 78(2), 453–508.