

Finding Mr. Schumpeter: An Empirical Study of Competition and Technology Adoption*

Jeffrey Macher[†]
Georgetown University

Nathan H. Miller[‡]
Georgetown University

Matthew Osborne[§]
University of Toronto

January 3, 2017

Abstract

We estimate the effect of competition on the adoption of a cost-reducing technology in the cement industry, using data that span 1953-2013. The new technology, the precalciner kiln, reduces fuel usage and hence fuel costs. We find adoption is more likely if the fuel cost savings are large, and less likely if there are many nearby competitors. We also find that competition damps the positive effect of cost savings. The results are consistent with a dynamic theoretical model in which competition can deprive firms of the scale necessary to recoup sunk adoption costs.

Keywords: technology, innovation, competition, portland cement
JEL classification: L1, L5, L6

*We thank Philippe Aghion, Jacob Cosman, Alberto Galasso, Richard Gilbert, Arik Levinsohn, Devesh Raval, John Rust, Rich Sweeney, Mihkel Tombak, Francis Vella, and seminar participants at Georgetown University, Harvard Business School, Harvard University, and the University of Toronto for helpful comments. We have benefited from conversations with Hendrick van Oss of the USGS and other industry participants. Conor Ryan provided research assistance.

[†]Georgetown University, McDonough School of Business, 37th and O Streets NW, Washington DC 20057. Email: jeffrey.macher@georgetown.edu.

[‡]Georgetown University, McDonough School of Business, 37th and O Streets NW, Washington DC 20057. Email: nathan.miller@georgetown.edu.

[§]University of Toronto, Rotman School of Management, 105 St. George Street, Toronto, ON, Canada M5S 3E6. Email: matthew.osborne@rotman.utoronto.ca.

1 Introduction

Joseph Schumpeter (1934, 1942) famously argues that large firms in concentrated markets invest more intensely in innovation. Empirical research on this hypothesis is challenging because innovation is difficult to measure and because the explanatory variables—market concentration and firm size—are endogenous economic outcomes. Further, the substantial reduced-form literature often relies on cross-industry comparisons that are difficult to interpret because the theoretical mechanisms that dominate in one industry may be unimportant in others.¹ Interest in the subject remains strong, however, and recent research makes progress by using dynamic structural models to simulate the effect of competition in specific settings (e.g., Goettler and Gordon (2011); Igami (2015); Igami and Uetake (2016)).

We contribute to this literature with a study of technology adoption in the portland cement industry. We first develop a dynamic model of technology adoption to understand the economic mechanisms at play. We then estimate the determinants of adoption without imposing (much) structure on the model, which allows us to reach conclusions about technology adoption that rely on the underlying empirical variation to the greatest extent possible. This approach is made possible by the data at our disposal, which span the period 1953-2013 and include hundreds of older kilns that could be replaced with fuel efficient precalciner kilns. The first precalciner kiln is installed in 1974, and precalciner kilns account for 74 percent of industry capacity by the end of the sample. Most older kilns that are not replaced are shut down instead. The analysis makes use of time-series variation in fossil fuel prices, as well as time-series and cross-sectional variation in competition and demand.

The theoretical objective of the paper is to understand adoption for a particular class of cost-reducing technologies. This class includes technologies for which adoption is (i) non-drastic because equilibrium prices remain above the marginal cost of old technology; (ii) non-divisible because adoption costs do not scale with capacity; and (iii) non-exclusive because adoption by one firm does not preclude others from adopting. This technology class is most common in process-intensive industries (e.g., Abernathy and Utterback (1978)), and arguably captures the most salient characteristics of precalciner adoption. We analyze a simplified version of the dynamic oligopoly models that are used in theoretical and computational research to understand investment, entry, and exit decisions under uncertainty (e.g., Ericson and Pakes (1995); Doraszelski and Satterthwaite (2010)). We show that competi-

¹Aghion and Tirole (1994) refer to the Schumpeterian hypothesis as the second most tested relationship in industrial organization, following only the price-concentration relationship. Even the literature reviews are daunting (e.g., Kamien and Schwartz (1982); Baldwin and Scott (1987); Cohen and Levin (1989); Cohen (1995); Gilbert (2006); Cohen (2010)).

tion limits long run adoption at least weakly. The mechanism is simple: competition denies firms the scale necessary to recoup adoption costs. The presence of many competitors does create preemption incentives and, in some parameterizations, competition speeds short-term adoption even as it limits long run adoption.

The empirical analysis centers on multinomial probit regressions that characterize kiln adoption and shutdown. The regressions can be interpreted as implementing the first step in the standard two-step estimator for dynamic games (e.g., Bajari, Benkard and Levin (2007); Ryan (2012)). We focus exclusively on the first step because we seek to understand firm policies rather than conduct counterfactual simulations. The institutional details of the industry—market power is localized due to high transportation costs—allow us to measure competition based on the number of nearby competing plants. The main challenge for identification is that this competition measure could be positively correlated with the unobserved error term, such that estimation risks understating the extent to which competition deters adoption. We proceed using the control function approach of Rivers and Vuong (1988). The excluded instrument is a 20-year lag on the competition variable. This lag has power because kilns are long-lived, and it is valid provided that autocorrelation in the unobserved error terms is not too great. We find that using a 15-year lag instead produces similar results, which supports that the persistence of the error term dies out over longer time horizons.

The regression results indicate that plants facing greater competition are less likely to adopt precalciner technology and more likely to shutter their older kilns. These effects are statistically significant and robust across a number of alternative modeling assumptions. The mean elasticity of the adoption probability with respect to competition ranges between -1.45 and -2.16 in the baseline specifications. Interpreted through the lens of the theoretical model, the empirical analysis supports that Schumpeterian scale effects dominate preemption effects in the data. Supplementary checks do not find evidence of strong preemption effects: for example, whether a plant is nearby early adopters has little additional explanatory power on precalciner adoption and kiln shutdown. The other comparative statics of the theoretical model find robust support in the data: the regressions indicate that adoption increases with the fuel costs savings provided by precalciners (which depend on fossil fuel prices), the amount of nearby construction activity, and the average fuel costs of nearby competitors. These results reinforce the connections between the model and the data.

Our research has bearing on whether carbon taxes induce firms to adopt green technology. Recent empirical articles on the so-called “induced innovation” hypothesis typically find some margin of adjustment but do not address competition (e.g., Newell, Jaffe and Stavins (1999); Popp (2002); Linn (2008); Aghion et al (2012); Hanlon (2014)). Our theoretical

model indicates that firms are more responsive to carbon taxes if competition is weaker. While we do not observe carbon taxes in the data, fossil fuel prices vary considerably over 1973-2013 and can proxy for carbon taxes. Our regression results imply that monopolists are nearly five times more likely than the average firm to install a precalciner kiln in response to higher fuel costs.² Thus, our theoretical and empirical results again are consistent. The policy insight is that innovation subsidies could be an important complement to market-based regulation in competitive markets; a similar result is obtained from the endogenous growth model of Acemoglu, Akcigit, Hanley and Kerr (2016).

Our research also is relevant to merger review. Allegations that mergers damp innovation incentives appear with some frequency in the Complaints of the DOJ and FTC (Gilbert (2006)). It is often difficult for outside economists to evaluate the merits of these allegations, however, because the court documents typically do not elaborate on the theoretical mechanism by which market structure affects innovation incentives. Our results suggest a specific setting in which mergers could have pro-competitive effects on innovation by allowing firms to achieve the scale required to profitably recoup the fixed costs of investment. Such mergers enhance firms' abilities to appropriate the returns to innovation; in this sense, our research relates to the large literature on appropriability reviewed in Cohen (2010).

Our empirical results should extend to the specific class of technologies defined above in markets that are at least reasonably competitive. External validity outside this class is better evaluated based on the large theoretical literature on competition and innovation. For example, while market power can facilitate innovation due to Schumpeterian effects (e.g., Dasgupta and Stiglitz (1980)), the opposite result obtains if innovation cannibalizes monopoly profit (Arrow (1962)), deters entry (Gilbert and Newbery (1982)), or allows firms to escape competitive pressure (Aghion et al (2005)).³ Even within the class of non-drastring, non-divisible, and non-exclusive technologies, our theoretical results indicate competition can speed short-term adoption. The lack of support for preemption in the data may be due to the large number of competitors that the average cement plant faces. This characteristic distinguishes the cement industry from the monopolies and tight oligopolies for which there is empirical support for preemption (e.g., Genesove (1999); Vogt (2000); Dafny (2005); Schmidt-Dengler (2006); Ellison and Ellison (2011); Gil, Houde and Takahashi (2015); Fang

²Plants with many nearby competitors are more likely to shut down their older kilns in response to higher fuel costs. This suggests a complicated industry adjustment process that we do not seek to model. The dynamic game would incorporate a "war of attrition" as studied by Ghemaway and Nalebuff (1985), Fudenberg and Tirole (1986), and Takahashi (2015).

³The literature on competition and innovation is incredibly deep. We refer readers to the complementary literature reviews of Aghion and Griffith (2005) and Gilbert (2006).

(2016)).

Our research builds on the substantial literature on technology adoption. The earliest contributions study competitive environments (e.g., Griliches (1957)) and thus do not address the research questions examined here. Empirical support for the Schumpeterian prediction that firm size encourages technology adoption has been found in a number of settings, including various technologies in banking (Hannan and McDowell (1984); Akhavein, Frame and White (2005); Fuentes, Hernandez-Murillo and Llobet (2010)); coal-fired steam-electric generating technologies among electric utilities (Rose and Joskow (1990)); machine tools in engineering (Karshenas and Stoneman (1993)); and MRIs in hospitals (Schmidt-Dengler (2006)). The most notable counter-example is the basic oxygen furnace in the steel industry (e.g., Oster (1982)). Our empirical results do not provide direct evidence regarding firm size and technology adoption. However, indirect evidence is provided because the empirical results align tightly with the comparative statics of a theoretical model in which scale determines whether firms can recoup adoption costs.

Finally, the portland cement industry is well-studied in the literature due in part to the wealth of publicly available data. Most relevant is Fowlie, Reguant and Ryan (2016), which estimates a dynamic structural model and simulates the effects of carbon taxes. The model allows plants to make capacity and exit decisions, but does not incorporate cost-reducing technology adoption. Stage-game payoffs are determined by Nash-Cournot competition. The simulations indicate that carbon taxes induce exit and capacity reductions. Our empirical results support that higher fuel prices increase the propensity for older kilns to shut down, and show that technology adoption also is an important margin of adjustment. Chicu (2012) estimates a dynamic structural model based on data from 1949-1969. Simulations on plants in Arizona—a duopoly state—suggest that preemption spurs capacity investments.

The paper proceeds as follows. Section 2 develops the theoretical model. Section 3 provides institutional details on precalciner kilns and the portland cement industry. Section 4 details the econometric methodology and identification strategy. Section 5 defines the variables used in the empirical analysis and provides summary statistics. Section 6 describes the results of the regression analysis, and Section 7 concludes.

2 Theoretical Model

2.1 Framework, policies, and equilibrium

We develop and analyze a dynamic model of cost-reducing technology adoption, building on the methodologies of Ericson and Pakes (1995) and Doraszelski and Satterthwaite (2010). Marginal costs equal c_1 and c_0 with and without the technology, respectively, with $c_1 < c_0$. Adoption is irreversible. We consider markets with $i = 1, \dots, N$ firms. The number of firms is exogenously determined and serves to scale the degree of competition. It is possible to endogenize N by incorporating entry and exit, but we present the simpler model because the comparative statics are similar (though preemption matters somewhat more with exit). In each period, each firm i that has not adopted the technology receives a private draw on adoption costs, k_i , which is drawn from a continuous distribution $F(\cdot)$ with support $[\underline{k}, \bar{k}]$.⁴ These firms then decide whether to adopt. Lastly, all firms compete in a stage game that determines static profit. There is a single state variable, $L_t = 0, \dots, N - 1$, that governs adoption decisions: the number of competitors that already have adopted.

The equilibrium concept in the stage game is Nash-Cournot.⁵ Indexing the action of “adopt” as $x = 1$ and the action of “not adopt” as $x = 0$, static profit is given by $\pi(c_x, L_t; N)$ for $x \in \{0, 1\}$. Prices are determined according to a linear demand schedule, and we restrict attention to areas of the parameter space in which equilibrium quantities are positive for every firm. This setup conveys a number of desirable properties:

- (i) Adoption is non-drastic: $\pi(c_0, L_t; N) > 0$ for all L_t .
- (ii) Adoption increases stage game profit: $\pi(c_1, L_t; N) > \pi(c_0, L_t; N)$ for all L_t .
- (iii) Adoption reduces the stage game profit of competitors: $\pi(c_x, L_t; N) > \pi(c_x, L_t + 1; N)$ for all L_t and $x \in \{0, 1\}$.

⁴The support of the adoption cost distribution can be bounded or unbounded, i.e., it can be the case that $\underline{k} = -\infty$ or $\bar{k} = \infty$. A bounded distribution is theoretically attractive because if $\underline{k} \geq 0$ it rules out negative adoption costs. The empirical model uses an unbounded support, in the context of Probit regressions, which ensures that all observations can be rationalized.

⁵It follows that most direct antecedent to our theoretical model is Dasgupta and Stiglitz (1980), which studies cost-reducing technologies in Nash-Cournot equilibrium. The models share the mechanism that competition can deprive firms of the scale necessary to recoup adoption costs. The main difference is that the Dasgupta and Stiglitz model is a static game in which firms choose their level of cost-reducing investment, whereas our model considers a discrete cost reduction and incorporates dynamic elements such as preemption and the option value of deferring adoption. Adding dynamics produces a richer relationship between competition and technology adoption. Iskhakov, Rust and Schjerning (2015) study the dynamics of cost-reducing technology in Nash-Bertrand equilibrium.

(iv) The increase in stage game profit due to adoption decreases with the number of adopters: $\pi(c_1, L_t; N) - \pi(c_0, L_t; N) > \pi(c_1, L_t + 1; N) - \pi(c_0, L_t + 1; N)$ for all L_t .

These properties create an incentive for preemption. Consider that if firm i adopts the technology then the benefit of adoption is reduced for firm i 's competitors (property (iv)), and this can increase the profit of firm i (property (iii)).

We characterize firm behavior in a symmetric Markov-perfect equilibrium in pure strategies. The following assumption is standard (e.g., Doraszelski and Satterthwaite (2010)) and helps ensure the existence of equilibrium:

Assumption A1: (i) The number of firms is finite, $N < \infty$. (ii) Stage game profits are bounded, i.e., $|\pi| < \infty$ for $c \in \{c_0, c_1\}$, all values of $L_t < N$, all $N < \infty$. (iii) The distribution of adoption costs, $F(\cdot)$, has positive density over a connected support, and an expectation that exists. (iv) Firms discount future payoffs, that is, $\delta \in [0, 1)$. (v) Profit functions are symmetric, i.e., $\pi(c_x, L_t; N)$ is the same for all firms i .

Denote the value function heading into period t (i.e., before the cost draws are received) without adoption as $V_0(L_t; N)$ and with adoption as $V_1(L_t; N)$. Once the cost draws are received, each firm adopts the technology if and only if $v_1(L_t; N, k) > v_0(L_t; N)$, where $v_x(\cdot)$ denotes the expected discounted profit for action $x \in \{0, 1\}$. Evaluating this inequality requires that each firm integrate out over the actions of its competitors because the cost draws are privately observed. Given symmetry, the adoption probability of any single firm can be written as $P(L_t; N)$. Let the probabilities with which the state space transitions from L_t to $L_{t+1} = 0, 1, \dots, N-1$ be collected in the vector $\mathbf{P}_0(L_t; N)$. The first $L_t - 1$ elements of this vector equal zero because adoption is irreversible.⁶ It is helpful to write the profit and value functions in vector form. Let $\boldsymbol{\pi}(c_x, N) = (\pi(c_x, 0; N), \pi(c_x, 1; N), \dots, \pi(c_x, N-1; N))'$ and $\mathbf{V}_x(N) = (V_x(0; N), V_x(1; N), \dots, V_x(N-1; N))'$.

With this notation in hand, the expected discounted profit for each action has the expression:

$$\begin{aligned} \text{Adopt:} \quad v_1(L_t; N, k) &= \mathbf{P}_0(L_t; N)' (\boldsymbol{\pi}(c_1, N) + \delta \mathbf{V}_1(N)) - k, & (1) \\ \text{Not Adopt:} \quad v_0(L_t; N) &= \mathbf{P}_0(L_t; N)' (\boldsymbol{\pi}(c_0, N) + \delta \mathbf{V}_0(N)) \end{aligned}$$

The optimal policy takes the form of a cutoff rule: firm i adopts if $k < k^*(L_t; N)$, where $k^*(L_t; N)$ is the value of k such that $v_0(L_t; N) = v_1(L_t; N, k)$. In turn, this implies that the adoption probabilities are $P(L_t; N) = F(k^*(L_t; N))$.

⁶If $L_t = 0$ and $N = 2$ then $\mathbf{P}_0(0; 2) = (1 - P(0; 2), P(0; 2))$. If instead $L_t = 1$ then $\mathbf{P}_0(1; 2) = (0, 1)$.

The value function associated with adoption, $V_1(L_t; N)$, has an explicit solution because adoption is irreversible. Define the upper triangular matrix $\mathbf{\Pi}_0$ as follows:

$$\mathbf{\Pi}_0 = \begin{bmatrix} \mathbf{P}_0(0; N)' \\ \mathbf{P}_0(1; N)' \\ \vdots \\ \mathbf{P}_0(N-1; N) \end{bmatrix}. \quad (2)$$

This $(N \times N)$ matrix fully characterizes the state-space transition probabilities for any firm that does not adopt the technology. Once a firm adopts, however, it changes the adoption probabilities of its competitors in subsequent periods. Let the $(N \times N)$ matrix $\mathbf{\Pi}_1$ characterize the post-adoption transitions of competitors.⁷ Then the value functions associated with adoption are given by the vector:

$$\mathbf{V}_1(N) = \mathbf{\Pi}_0 (\mathbf{I} + \delta(\mathbf{I} - \delta\mathbf{\Pi}_1)^{-1}) \boldsymbol{\pi}(c_1, N). \quad (3)$$

The value function associated with not adopting is given by the following equation:

$$V_0(L_t; N) = \int_{\underline{k}}^{\bar{k}} \max\{v_0(L_t; N), v_1(L_t; N, k)\} dF(k). \quad (4)$$

Assumption A2: Define an industry state transition matrix, $\tilde{\mathbf{\Pi}}$, that characterizes the probabilistic changes in the total number of firms that have adopted the technology, \tilde{L}_t . The industry state transition matrix is continuous in each firm's adoption strategy, and the industry state \tilde{L}_t .

Under A1 and A2, Propositions 2 and 5 of Doraszelski and Satterthwaite (2010) guarantees the existence of a symmetric pure strategy Nash equilibrium. This result helps motivate our empirical model, which assumes that firms take identical actions given identical observables and unobservables. The dynamic game is simpler than that of Doraszelski and Satterthwaite because investment is discrete and exit is prohibited. These changes do not materially affect the proofs; it is possible to apply standard dynamic programming methods and Brouwer's fixed point theorem following the same arguments.⁸

⁷The matrix $\mathbf{\Pi}_1$ is composed of stacked vectors of post-adoption transition probabilities $\mathbf{P}_1(L_t; N)$. Because an adoption changes all subsequent adoption probabilities, $\mathbf{P}_1(L_t; N)$ is different than $\mathbf{P}_0(L_t; N)$. For example, if $N = 2$ then $\mathbf{P}_1(0; N) = (1 - P(1), P(1))$, but $\mathbf{P}_1(1; N) = (0, 1)$.

⁸The key component is the assumption of continuous private shocks, which implies that a firm adopts if it receives a draw below $k < k^*(L_t; N)$. Without this, the existence of an equilibrium is not guaranteed

2.2 Preemption and the benefits of adoption

In this section, we derive the value to a firm of adopting the new technology, in order to understand the different forces that drive that decision. Define the benefit of adoption as $b(L_t; N, k) = v_1(L_t; N, k) - v_0(L_t; N)$. Stacking across states yields the vector $\mathbf{b}(N, k) = (b(0; N, k), b(1; N, k), \dots, b(N-1; N, k))'$. Based on the value functions shown in equation (3) and (4), it is possible to construct a set of intermediary matrices such that:

$$\begin{aligned} \mathbf{b}(N, k) = & \underbrace{\mathbf{\Pi}_0(\boldsymbol{\pi}_1 - \boldsymbol{\pi}_0) - \boldsymbol{\kappa}}_{\text{Term 1}} + \underbrace{\delta \mathbf{\Pi}_0 \mathbf{A}(\mathbf{\Pi}_0, \mathbf{\Pi}_1) \boldsymbol{\pi}_1}_{\text{Term 2}} - \underbrace{\delta \mathbf{\Pi}_0 \mathbf{B}(\mathbf{\Pi}_0) \boldsymbol{\pi}_0}_{\text{Term 3}} \\ & - \underbrace{\delta \mathbf{\Pi}_0 \mathbf{B}(\mathbf{\Pi}_0) \mathbf{D} [\mathbf{\Pi}_0(\boldsymbol{\pi}_1 - \boldsymbol{\pi}_0) - \boldsymbol{\kappa} + \delta \mathbf{\Pi}_0 \mathbf{A}(\mathbf{\Pi}_0, \mathbf{\Pi}_1) \boldsymbol{\pi}_1]}_{\text{Term 4}} \end{aligned} \quad (5)$$

The vector $\boldsymbol{\kappa}$ gives the expected adoption costs; the L^{th} element of that vector equals $E(k|k < k^*(L-1))$. The matrix \mathbf{D} summarizes adoption probabilities; it is diagonal and has $F(k^*(L))$ as the $(L+1)^{\text{st}}$ diagonal element. The matrices $\mathbf{A}(\mathbf{\Pi}_0, \mathbf{\Pi}_1)$ and $\mathbf{B}(\mathbf{\Pi}_0)$ serve to discount future profit streams and are provided in Appendix B.1.

The first term in equation (5) is the expected increase in static profit due to adoption that would be realized in the subsequent stage-game, less the observed adoption cost. The second term is the expected discounted stream of future profit associated with adoption. It incorporates the post-adoption transition matrix $\mathbf{\Pi}_1$, which captures how competitors respond to adoption. This is how preemption affects decisions: adoption reduces the probability that competitors adopt, and thereby increases the profit flows that arise with adoption. The third term is the expected discounted stream of future profit associated with not adopting. It incorporates only the explicit profit, i.e., the profit earned before the firm adopts in some future period. The fourth term captures the option value of waiting to adopt: if a firm does not adopt today, it might find it appropriate to do so in the future.

Figure 1 considers two simple numerical examples to illustrate the role of preemption. The left column shows the optimal policy functions and the right column shows the expected time path of adoption. The top and bottom rows correspond to the two different parameterizations. Each panel features results from the full model (“with preemption”) and an alternative model that eliminates the preemption incentive (“without preemption”). The latter model is solved by substituting $\mathbf{\Pi}_0$ for $\mathbf{\Pi}_1$ in the value functions, so that firms do not consider that adoption changes the subsequent adoption probabilities of competitors.

without admitting mixed strategies. The issue that can arise is similar to that of Ericson and Pakes (1995) in the context of entry or exit, because adoption also is a discrete action. Allowing for private shocks means that an individual firm essentially treats its rivals as mixing over upgrading and not upgrading.

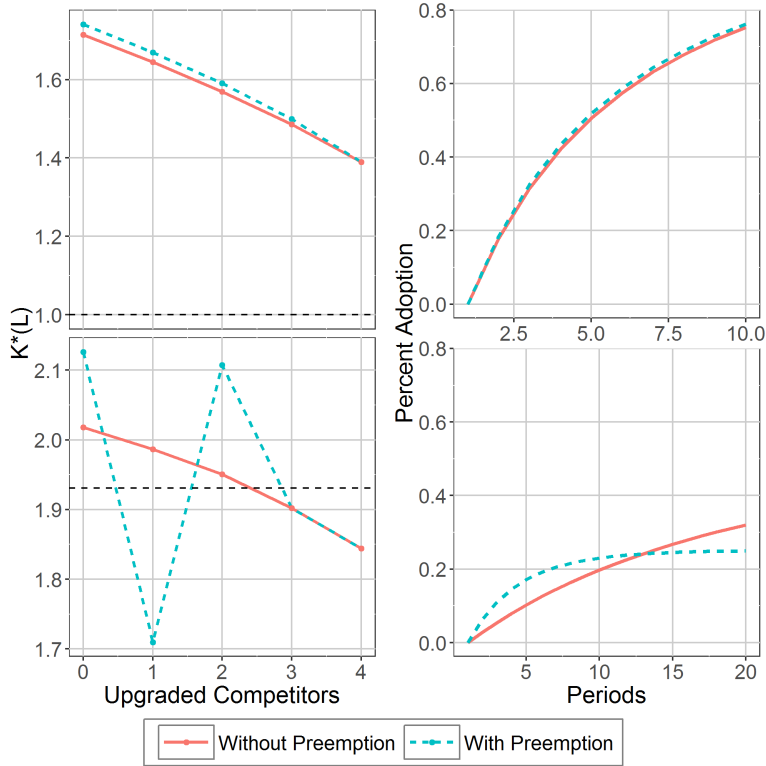


Figure 1: The Role of Preemption

Notes: Each row summarizes the model under a different parameterization. The left panels provide the equilibrium policy functions as a function of the state space (L_t). The right panels show expected time paths in terms of the fraction of the industry that has adopted (L_t/N). The horizontal black dashed lines in the left panels show \underline{k} , the lower bound of adoption costs. Appendix Table C.1 provides the parameterizations.

In the top row, preemption shifts the equilibrium cutoff rule such that firms adopt for higher values of k , unless no other competitors would adopt the technology, in which case the models are identical. The industry adopts the technology faster with preemption—though the difference is quite small—but converges to the same long run equilibrium. This pattern holds qualitatively over the many parameterizations that we have investigated. Preemption has larger effects in the bottom row. Not every firm adopts in the long run equilibrium and this induces something of a “race” to adopt. The cutoff level of k for the first adopter (i.e., with $L_t = 0$) is higher with preemption and this leads to noticeable faster adoption.⁹

There is at least one special case of the model in which preemption plays little role. If

⁹There is no second adopter in the parameterization with preemption if the state $L_t = 1$ is realized because $k^*(1; N) < \underline{k}$. However, if two firms adopt simultaneously at state $L_t = 0$ then the industry moves directly to the $L_t = 2$ state and a third firm adopts.

adoption probabilities are small then the benefits of adoption can be approximated as:

$$\mathbf{b}(N) \approx \frac{1}{1-\delta} (\boldsymbol{\pi}_1 - \boldsymbol{\pi}_0) - k, \quad (6)$$

Both the preemption and option value terms in equation (5) become small because future adoption is unlikely, and the equilibrium cutoff rule simplifies to a comparison of stage game profit with and without adoption. This special case helps motivate the next subsection, in which we explore the comparative statics of the Nash-Cournot model and develop that firm scale is a fundamental determinant of technology adoption. It also has bearing on the empirical exercise because the unconditional probability of precalciner adoption is less than two percent in a given year.

2.3 Comparative statics and firm size

In this section we focus on the adoption decision of a firm in a situation in which its competitors are unlikely to adopt. We model this decision using a “focal firm” setting, which means that the focal firm does not integrate out the upgrade probabilities of its rivals. Letting the inverse demand curve be given by $P(Q) = a - Q$ where $Q = \sum_{j=1}^N q_j$, the equilibrium markups and quantities of the focal firm facing the adoption decision are given by:

$$q^*(c_x, L; N) = P_x^*(L; N) - c_x = \frac{a - c_x + N(\bar{c} - c_x)}{(N + 1)} \quad (7)$$

where again $x \in \{0, 1\}$ denotes the action of the focal firm, and \bar{c} captures the average cost of all firms (i.e., $\bar{c} = \frac{1}{N}[Lc_1 + (N - L - 1)c_0 + c_x]$). The profit of the focal firm is $\pi(c_x, L; N) = (q^*(c_x, L; N))^2$ due to the equality of equilibrium markups and quantities.¹⁰

Following equation (6), with low adoption probabilities the benefit of adoption can be approximated as:

$$\begin{aligned} b(L; N) &\approx \frac{1}{1-\delta} (\pi(c_1, L; N) - \pi(c_0, L; N)) - k \\ &= \frac{1}{1-\delta} \frac{2N}{(N+1)} q^*(\hat{c}, L; N) \Delta c - k \end{aligned} \quad (8)$$

We refer to the second line as the “static benefit” of adoption because it does not incorporate

¹⁰We refer readers to Shapiro (1989) for a more general discussion of the Nash-Cournot model, including conditions for the existence and uniqueness of equilibrium with nonlinear demand. We employ a unit slope normalization. This is without loss of generality and all results are robust to demand curve rotations.

preemption or the option value. It can be derived in a few lines of algebra starting with equilibrium markups and quantities, and includes both the magnitude of the cost savings ($\Delta c = c_0 - c_1$) and the midpoint cost level ($\widehat{c} = \frac{1}{2}(c_0 + c_1)$). The following comparative statics are then straight-forward:

Result 1: *The static benefit of adoption (i) increases in the cost savings, Δc ; (ii) decreases in the initial cost level, c_0 ; (iii) increases with industry average costs, \bar{c} ; and (iv) increases with the inverse demand intercept, a . Further, profit decreases with costs and the number of firms, but increases with the industry average costs and the inverse demand intercept.*

The effect of competitors (i.e., N) on the static benefit is difficult to sign with a tractable expression because adding a competitor j with costs $c_j \in \{c_0, c_1\}$ changes industry average costs. Consider instead the effect of competitors that have marginal costs $c_j = \bar{c}$. (Such competitors do not exist in the full model, but this approach nonetheless conveys useful intuition.) An additional competitor decreases the benefit that the focal firm receives from adoption if the following condition holds:

$$\left(\frac{\bar{c} - \widehat{c}}{\widehat{c}}\right) < \left(\frac{a - \widehat{c}}{\widehat{c}}\right) \left[\frac{(N + 1)^3 - N(N + 2)^2}{N^2(N + 2)^2 - (N + 1)^4}\right] \quad (9)$$

We sketch the derivation in Appendix B. The term in brackets equals 0.29 if $N = 2$ and converges to 0.50 as N grows large. The condition holds trivially if $L > N/2$ because then the LHS is negative and is exceeded by the RHS (which is positive). In earlier stages of industry adoption, the condition holds provided that the cost reduction of the technology is not too great relative to the total surplus created by the industry, which seems plausible in many settings.¹¹ This leads to a second set of comparative statics:

Result 2: *Under condition (9), increasing the number of competitors (i) decreases the static benefit of adoption; and (ii) decreases the (positive) effect of Δc on the static benefit.*

We pause here to consider that the cost savings of the technology increase the static benefit of adoption, but that this effect diminishes with the number of firms. This is relevant for the induced innovation hypothesis because it suggests that competition can mitigate the

¹¹The condition almost surely holds in our empirical application. Because precalciner technology reduces fuel costs by about 30 percent, the maximum value the LHS could take in our application is around $(1 - 0.85)/0.85 = 0.176$. The empirical results of Ganapati, Shapiro and Walker (2016) can be manipulated to obtain $(p - c)/c = 1.50$ for the cement industry. This provides a lower bound to $(a - \widehat{c})/\widehat{c}$. The Ganapati, Shapiro and Walker results use data from the Census of Manufacturers. See Table 2 of the May 2016 draft, which indicates prices of 0.05 and costs of 0.02 (in thousands of 1987 dollars per cubic yard).

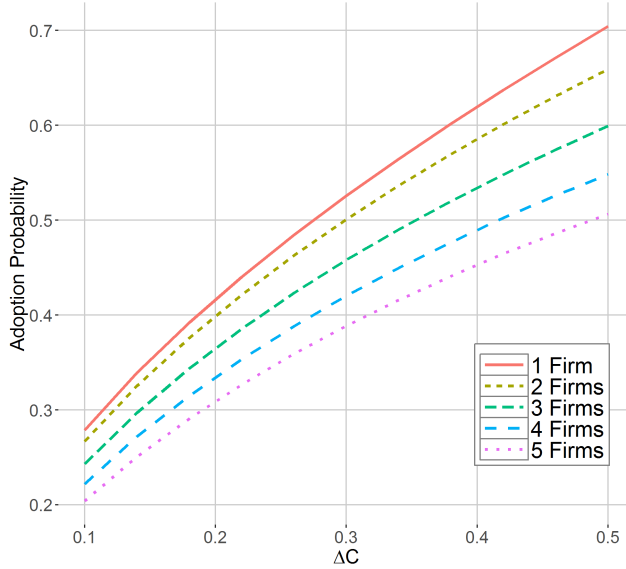


Figure 2: Induced Innovation in the Theoretical Model

Notes: The vertical axis is the adoption probability given $L_t = 0$ (i.e., $F(k^*(0; N))$). The horizontal axis is Δc , the magnitude of the cost savings. The different lines show the relationship for $N = 1, \dots, 5$. Appendix Table C.1 provides the parameterization.

responsiveness of firms to some innovation stimulus. One policy-relevant example is carbon taxes, which affect adoption incentives by growing the cost difference between “green” and “dirty” technologies. Figure 2 shows that the comparative static holds in the fully dynamic version of the model, in which we do not restrict firms to believe its rivals do not upgrade. The vertical axis is the probability of adoption given $L_t = 0$. The horizontal axis show different levels of Δc . The relationship between adoption and Δc is provided for $N = 1, \dots, 5$. Adoption probabilities increase for larger Δc , consistent with part (i) of Result 1, but the increase is less pronounced with large N , consistent with part (ii) of Result 2.

To develop the mechanisms underlying the comparative statics, we reconsider the static adoption benefits of equation (8) using a first order Taylor Series expansion:

$$\pi(c_1, L; N) - \pi(c_0, L; N) \approx - \left. \frac{\partial \pi(c, L; N)}{\partial c} \right|_{c=\hat{c}} \Delta c \quad (10)$$

where again $\Delta c = c_0 - c_1$ and $\hat{c} = \frac{1}{2}(c_0 + c_1)$. This expansion holds with equality if demand is linear but only differentiability is necessary for a more general analysis. An application of

the envelope theorem yields the derivative of the profit function:

$$-\frac{\partial \pi(c, L_t; N)}{\partial c} = \underbrace{q_i^*(\hat{c}, L; N)}_{\text{Scale Effect}} - \underbrace{q_i^*(\hat{c}, L; N) \sum_{k \neq i} \frac{\partial P^*}{\partial q_k} \frac{\partial q_k}{\partial q_i} \frac{\partial q_i^*}{\partial c}}_{\text{Strategic Effect}} \quad (11)$$

where we use i to index the focal firm. The scale effect captures the standard intuition that reductions in marginal cost are more profitable if scaled over many units of output (Arrow (1962)). The magnitude of the scale effect decreases with the number of competitors, and this creates Schumpeterian dynamics in the full model: competition can deny firms the scale necessary to recoup adoption costs. By contrast, equilibrium output increases with demand and industry average costs, and this allows firms to more easily recoup adoption costs. The strategic effect captures that adoption induces competitors to produce less. Its magnitude can increase with the number of competitors. With linear demand, the strategic effect simplifies to $\left(\frac{N-1}{N+1}\right) q^*(\hat{c}, L; N)$, and condition (9) determines the net effect of competition on the profit derivative.

2.4 Competition and adoption in the dynamic game

Figure 3 provides graphical intuition about the relationship between competition and technology adoption in the fully dynamic model which allows for preemption. The rows corresponds to different parameterizations. The left column provides the equilibrium policy functions as a function of the state space (L_t). The horizontal black dashed lines show \underline{k} , the lower bound of the adoption costs. The right column shows the expected time path of \tilde{L}_t/N , the fraction of the industry that adopts.

The graph of the policy function in the top row exhibits several notable features. First, every firm adopts eventually because the equilibrium cutoff levels exceed the lower bound of the adoption costs (i.e., $k^*(L_t) > \underline{k}$ for all L_t and N). Second, the policy functions slope down, for a given N , which indicates that adoption is more likely if fewer competitors have already adopted. This result happens because with large L_t preemption incentives are weaker and the stage game benefits are smaller. Third, adding competitors to the model shifts the policy functions down so that adoption is less likely at any level of L_t . The graph of the time path shows how these features play out. Adoption happens faster if there are fewer firms, but regardless an absorbing state is eventually reached in which all firms adopt.

The second row differs because the equilibrium cutoff levels do not always exceed the lower bound of the adoption cost distribution. Specifically, only three firms adopt if $N = 4$

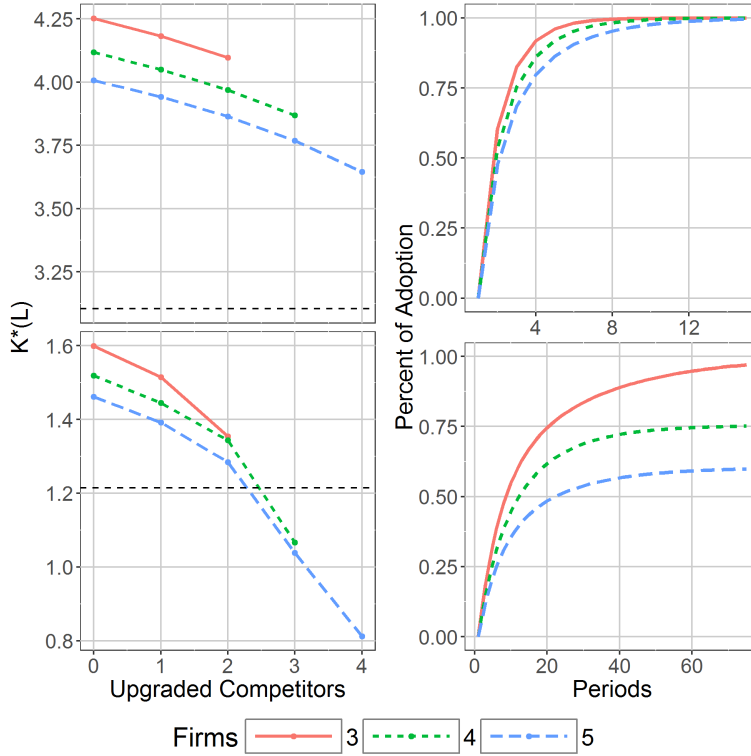


Figure 3: Competition and Technology Adoption

Notes: Each row summarizes the model under a different parameterization. The left column provides the equilibrium policy functions as a function of the state space (L_t). The right column shows expected time paths in terms of the fraction of the industry that has adopted (\tilde{L}_t/N). The solid red, dotted green, and dashed blue lines correspond to $N = 3$, $N = 4$, and $N = 5$, respectively. The horizontal black dashed lines that appear in the left column show \underline{k} , the lower bound of the adoption cost distribution. Appendix Table C.1 provides the parameterizations.

or $N = 5$. The expected time paths show that competition both slows adoption and reduces adoption in the absorbing state. The long run effect of competition arises because having many competitors denies some firms the scale necessary to recoup adoption costs; preemption does not matter for the last adopter, so it does not affect long run adoption.

Figure 3 supports a conjecture that competition limits long run adoption at least weakly. Indeed, the conjecture holds for a special class of absorbing states in which reversing the adoption of a single firm would generate further adoption at a cost draw of \underline{k} .¹² Because time is discrete, adoption can overshoot this class of absorbing states if multiple firms receive advantageous (private) cost draws. This is unlikely, however, because adoption

¹²Formally, states in the special class are characterized by a number of adopters \tilde{L}^* such that the $(\tilde{L}^* + 1, \tilde{L}^* + 1)$ element of the industry transition matrix $\tilde{\Pi}$ equals one, and the $(\tilde{L}^*, \tilde{L}^* + 1)$ element of $\tilde{\Pi}$ is positive but less than one.

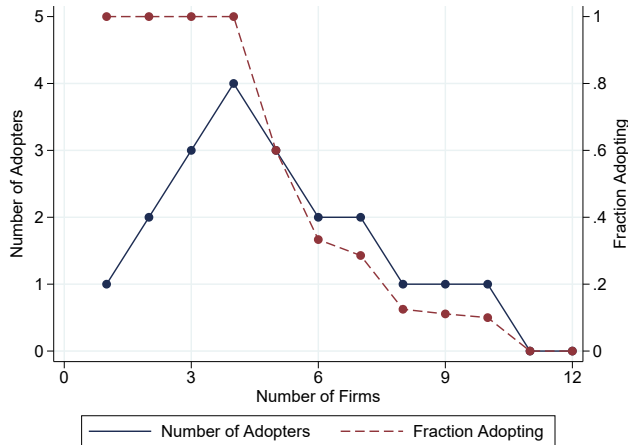


Figure 4: Technology Adoption in the Absorbing States

probabilities tend to be small near the absorbing state. The proposition below characterizes how the number of adopters in the class of absorbing states, L^* , changes with the number of competitors.

Proposition 1: *If $\underline{k} \leq 0$ then $L^* = N$. If instead $\underline{k} > 0$ there exists some n_1 such that if $N > n_1$ then $L^* < N$ and L^* weakly decreases in N . Further, if $\underline{k} > 0$ and $(1 - \delta)\underline{k} > (\Delta c)^2$ then there exists some $n_2 > n_1$ such that if $N > n_2$ then $L^* = 0$.*

The proof of the proposition is in Appendix B. We show that it is possible to characterize adoption in the special class of absorbing states analytically for any N , which extends inference beyond the limits of numerical solution. To reinforce intuition, Figure 4 plots the number and fraction of adopters under one specific parameterization. For $N \leq 4$, the number of adopters grows with the number of firms because all firms find it profitable to adopt. The number of adopters shrinks for $N > 4$, and equals zero for $N \geq 11$. The fraction of firms that adopt begins at one, and then falls to zero over $N \in [4, 11]$. Competition eventually becomes sufficient to deny at least one firm the scale necessary to recoup adoption costs, and thereafter long run adoption decreases in the number of competitors.

The role of competition is more complicated in the early periods because preemption becomes relatively more important. Figure 5 summarizes results from another parameterization to make this point. As shown, competition can speed the pace of adoption in the short-term even as it limits adoption in the absorbing states. This is due to the preemption incentives, which can be particularly strong if not all firms eventually adopt. It also is readily transparent that an inverted-U operates in the initial state ($L_t = 0$) because adoption is more

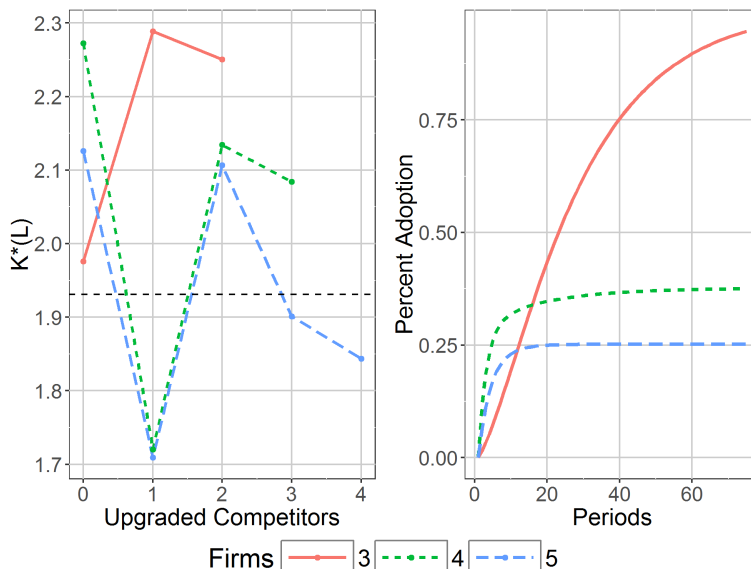


Figure 5: Competition and Technology Adoption

Notes: The left panel provides the equilibrium policy functions as a function of the state space (L_t). The right panel shows expected time paths in terms of the fraction of the industry that has adopted (\tilde{L}_t/N). The solid red, dotted green, and dashed blue lines correspond to $N = 3$, $N = 4$, and $N = 5$, respectively. The horizontal black dashed lines in the left panel show \underline{k} , the lower bound of adoption costs. Appendix Table C.1 provides the parameterization.

likely with $N = 4$ than with $N = 3$ or $N = 5$. Indeed, it is a simple corollary to Proposition 1 that an inverted-U exists whenever $k^*(L_t; N_2) > k^*(L_t; N_1)$ for $N_2 > N_1$, because there is always an $N_3 > N_2$ such that $k^*(L_t; N_2) > k^*(L_t; N_3)$. The inverted-U appears only in the early periods, however, because Proposition 1 dictates a monotonic negative relationship between competition and adoption in the absorbing states.

3 Empirical Setting

3.1 The portland cement industry

We examine the adoption of precalciner technology in the portland cement industry over 1973-2013. Portland cement is a finely ground dust that forms concrete when mixed with water and coarse aggregates such as sand and stone. Concrete, in turn, is an essential input to many construction and transportation projects. The production of cement involves feeding limestone and other raw materials into rotary kilns that reach peak temperatures of 1400-1450° Celsius. Fuel costs account for a sizable portion of overall revenues.

Plants equipped with precalciner technology preheat the raw materials using the ex-

haust gases of the kiln combined with heat from a supplementary combustion chamber, which reduces production energy requirements by 25-35 percent by allowing an important chemical reaction (calcination) to begin before raw materials enter the kiln. This reduces the requisite kiln length and requires a complete plant retrofit. Cement producers outsource kiln design to one of several industrial architecture firms with expertise in cement. The physical component is not especially demanding—many industrial construction firms can manage the steel plates, refractory linings, and duct work—but total design and installation costs are large. To provide an order of magnitude, publicly-available estimates place the total cost of building a modern cement plant around \$800 million.¹³

Table 1 tracks precalciner kiln adoption over time. In 1973, nearly all plants used inefficient wet and long dry kilns. A small number of plants utilized preheater technology, which recycles exhaust gases without a supplementary combustion chamber, but no plant used precalciner technology. Over the ensuing four decades, the number of wet kilns decreased from 249 to 19 and the number of long dry kilns decreased from 157 to 26. Shuttered kilns typically remain on site because they are costly to relocate, but most of the supporting equipment can be repurposed profitably. By the final year of data 2013, there are 66 precalciner kilns in operation and these account for 74 percent of industry capacity.

Table 2 provides the average fuel costs among kilns in each technology class, again at five-year intervals over the sample period. These costs are obtained based on kiln efficiency and the price/mBtu of the primary fossil fuel used. The changes within kiln technology classes over time are driven primarily by exogenous fluctuations in natural gas and coal prices, which provides a key source of variation that we exploit in the estimation. This feature of the data can be interpreted further as providing a natural experiment as to how firms would respond to carbon taxes, which change the price of fossil fuels.

Table 2 also provides the fuel costs of the “frontier technology,” which we define as a precalciner kiln that burns the most affordable fuel. The difference between a kiln’s fuel cost and that of the frontier technology – a measure of the fuel cost savings available from precalciner adoption – is an empirical analog to the Δc term in the motivating theory. Fuel cost savings tend to be large when fossil fuel prices (and thus fuel costs) are high.¹⁴ The

¹³CEMBUREAU, the European cement association, places construction costs for a one million metric tonne plant at around three years of revenue, and estimates annual total costs of around \$200 million. A study by The Carbon War Room (2011), an environmental action group, places profit margins at 33 percent given a per-tonne price of \$100. Putting these facts together, our \$800 million number is calculated as $\$200 \times 1.33 \times 3 = \$798 \approx \$800$. For the CEMBUREAU estimate, see <http://www.cembureau.be/about-cement/cement-industry-main-characteristics>

¹⁴There is a well known analogy in the automobile industry: the driving cost of vehicles with low miles-per-gallon (MPG) is more sensitive to the gasoline price than that of high MPG vehicles, and automobile

Table 1: The Portland Cement Industry over 1973-2013

Year	Wet Kilns	Long Dry Kilns	Dry with Preheater	Dry with Precalciner	Total Kilns	Total Plants	Total Capacity
1973	249	157	23	0	429	159	76.67
1978	201	111	42	2	356	151	79.85
1983	121	90	36	24	271	132	79.79
1988	96	70	35	26	227	116	75.47
1993	72	65	38	27	202	107	74.50
1998	67	63	34	31	195	106	76.79
2003	53	49	38	45	185	106	90.88
2008	45	31	32	56	164	103	96.00
2013	19	26	29	66	140	95	98.45

Notes: The table shows data at five-year snapshots spanning 1973-2013. Kiln counts are provided separately for each of the four production technologies: wet kiln, long dry kilns, dry kilns with preheaters, and dry kilns with precalciners. Total capacity is in millions of metric tonnes. The data are for the contiguous U.S. and are obtained from the *PCA Plant Information Survey*.

final column of the table provides the national average price of portland cement: depending on the year and kiln technology, fuel costs account for between 8 and 33 percent of revenues. Two recent papers estimate that pass-through of fuel costs to price in the cement industry exceeds unity (Miller, Osborne and Sheu (2015); Ganapati, Shapiro and Walker (2016)).

Cement is typically transported by truck to ready-mix concrete plants and large construction sites, and these associated costs generally account for a sizable portion of purchasers' total expenditures. The academic literature often models the industry as a number of distinct local markets (e.g., Ryan (2012); Fowlie, Reguant and Ryan (2016)). Figure 6 provides a map of the cement plants in operation as of 2010. Some geographic areas (e.g., southern California) have many plants, while others areas (e.g., South Dakota) have only a single nearby plant. These differences provide useful cross-sectional variation.

As cement is used in construction projects, demand is highly procyclical. Figure 7 graphs total production and consumption in the United States over 1973-2013. When macroeconomic conditions are favorable, consumption tends to outstrip production due to domestic capacity constraints; imports make up the differential. The technology by which cement can be shipped via transoceanic freighter at low cost and imported was developed in the late 1970s, which explains the tight connection between consumption and production in the earliest years of the sample. U.S. cement exports are negligible. Finally, cement cannot be stored for any meaningful period of time, because the product gradually absorbs moisture prices adjust accordingly (e.g., Busse, Knittel and Zettelmeyer (2013); Langer and Miller (2013)).

Table 2: Fuel Costs per Metric Tonne of Cement

Year	Wet Kilns	Long Dry Kilns	Dry with Preheater	Dry with Precalciner	Frontier Technology	Average Price
1973	18.99	16.41	13.30	.	10.33	85.59
1978	36.42	31.13	24.56	23.35	22.41	110.25
1983	28.84	23.63	18.06	16.78	16.78	94.41
1988	19.81	15.91	13.28	12.41	11.75	79.78
1993	15.35	12.66	9.86	9.77	9.44	77.97
1998	13.50	11.24	8.75	8.39	8.39	98.13
2003	12.94	11.26	8.76	8.40	8.40	87.53
2008	22.81	19.85	15.45	14.81	14.81	105.55
2013	25.70	22.36	17.40	16.83	16.68	89.93

Notes: The table provides average fuel costs by kiln technology, the hypothetical fuel costs of a kiln with “frontier technology” defined as a precalciner kiln that burns the most affordable fuel, and the national average price of portland cement. Data are shown at five-year snapshots spanning 1973-2013. Fuel costs are based on authors’ calculations as detailed in Appendix A. Prices are obtained from the USGS *Minerals Yearbook*. All statistics are in real 2010 dollars per metric tonne of cement output.

in the air which eventually renders it unusable.

3.2 Data sources

We draw on several data sources to construct a panel of kiln-year observations that span the contiguous United States over 1973-2013. This sample period is determined by the Portland Cement Association’s (PCA) *Plant Information Survey* (PIS), which is published annually over 1973-2003, semi-annually over 2004-2010, and then again in 2013. The PIS provides an end-of-year snapshot of the industry that includes the location, owner, and primary fuel of each cement plant in the U.S. and Canada, as well as the age, capacity and technology class of each kiln. We impute values in missing years by using data from preceding and following years, as well as by using information in the *Minerals Yearbook* of the United States Geological Survey (USGS), which summarizes an annual cement plant census. We combine the PIS kiln data with supplementary data that contain kiln locations over 1949-1973. These data were constructed by backcasting the 1973 PIS using information culled from the trade publication *Pit and Quarry*, occasionally printed *Pit and Quarry* maps of the industry, and the *American Cement Directory*. We refer readers to Chicu (2012) for details.¹⁵ The supplementary dataset is useful because it allows us to construct lagged competition

¹⁵We thank Mark Chicu for making these data available.

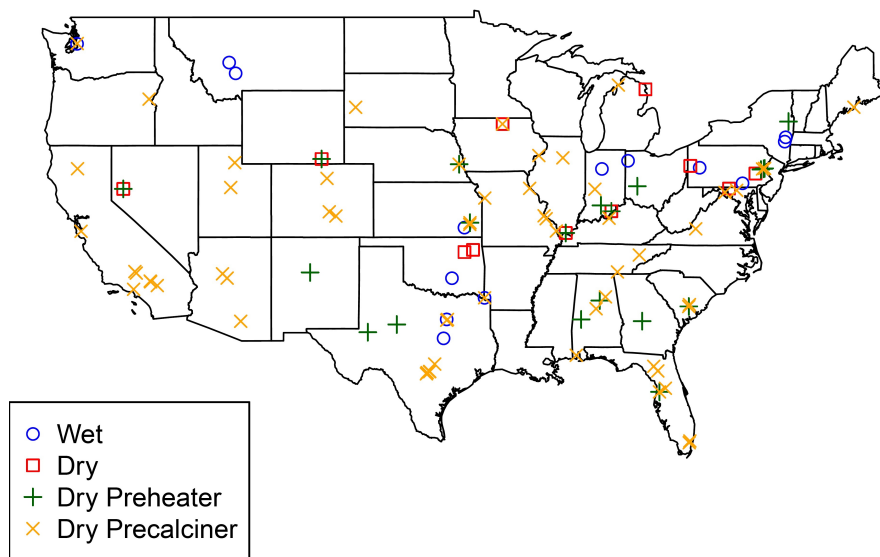


Figure 6: Portland Cement Plants in the Contiguous United States, in 2010

measures without discarding the earlier years of the PIS sample.

We calculate the fuel costs of production based on kiln efficiency and fossil fuel prices, using the PCA’s *U.S. and Canadian Portland Cement Labor-Energy Input Survey* to measure production energy requirements. This survey is published intermittently, and we use the 1974-1979, 1990, 2000, and 2010 versions. We obtain the average prices of coal, natural gas, and distillate fuel oil for the industrial sector from the State Energy Database System (SEDS) of the Energy Information Agency (EIA). We use fossil fuel prices at the national level because they are more predictive of cement prices (Miller, Osborne and Sheu (2015)), probably due to the measurement error associated with imputing withheld state-level data. We obtain retail gasoline prices from the EIA’s *Monthly Energy Review*.¹⁶ We use county-level data on construction employment and building permits from the Census Bureau to account for demand-side fluctuations.¹⁷ Construction employment is part of the County Business Patterns data. We use NAICS Code 23 and (for earlier years) SIC Code 15. The data for 1986-2010 are available online.¹⁸ The data for 1973-1985 are obtained from the

¹⁶The gasoline prices include federal and sales taxes, and are for regular leaded gasoline until 1990 and regular unleaded gasoline thereafter. See <http://energy.gov/eere/vehicles/fact-915-march-7-2016-average-historical-annual-gasoline-pump-price-1929-2015>, last accessed April 25, 2016.

¹⁷For both the construction employment and building permits, it is necessary to impute a small number of missing values. We calculate the average percentage difference between the observed data of each county and the corresponding state data, and use that together with the state data to fill in the missing values.

¹⁸See <http://www.census.gov/econ/cbp/download/>, last accessed April 16, 2014.

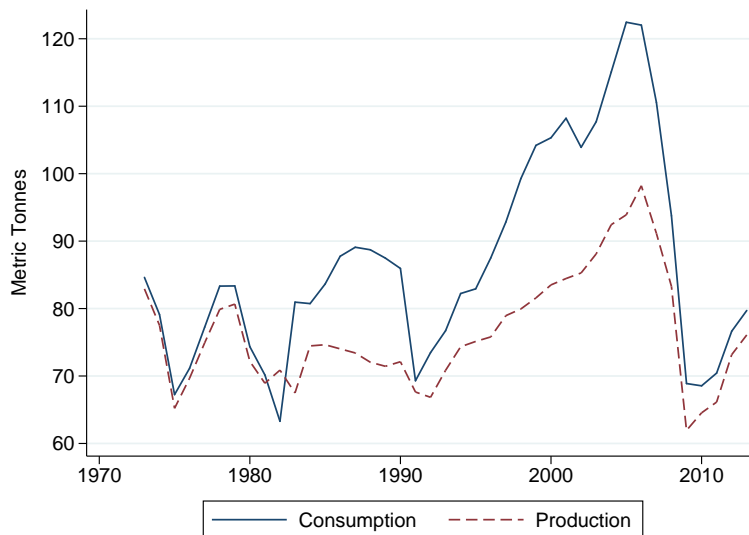


Figure 7: Consumption and Production in the United States, 1973-2013

University of Michigan Data Warehouse. The building permits data are maintained online by the U.S. Department of Housing and Urban Development.¹⁹ Finally, data on cement prices, consumption, and production reported in the previous subsection are obtained from the USGS *Minerals Yearbook*. USGS does not provide firm-level or plant-level data.

4 Empirical Model

4.1 Policy functions

Our empirical objective is to characterize the policy functions that govern technology adoption and kiln shutdown. We use multinomial probit regressions to implement the first step from the standard two-step estimator for dynamic games developed in Bajari, Benkard and Levin (2007) and applied in research such as Ryan (2012). The second step, which uses forward simulation to recover dynamic structural parameters, is unnecessary because we do not conduct counterfactual simulations.

To formalize our approach, consider that profit in the stage game of the theoretical model depends on cost, industry average cost, demand, and the number of firms. Let the empirical analog be $\pi(c_{it}^x, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta)$, where the subscripts i and t identify the kiln and year, respectively. The first four arguments are defined as in the theoretical model (for

¹⁹See <http://socds.huduser.org/permits/>, last accessed April 16, 2014.

$x = \{0, 1\}$), w_{it} is a vector of controls, and θ is a vector of parameters. Define the empirical “benefit of adoption” for a producer with an old kiln as:

$$b(\Delta c_{it}, c_{it}^0, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta) = \pi(c_{it}^1, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta) - \pi(c_{it}^0, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta) - k_{it}$$

With this notation in hand, the technology choice of producers can be framed as a maximization problem in which the maximand is:

$$\Pi_{it} = \begin{cases} b(\Delta c_{it}, c_{it}^0, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta) + u_{it}^A & \text{if adopt} \\ u_{it}^0 & \text{if maintain} \\ -\pi(c_{it}^0, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta) + u_{it}^S & \text{if shut down} \end{cases} \quad (12)$$

where $(u_{it}^A, u_{it}^0, u_{it}^S)$ are the stochastic shocks. We parameterize the functions using linear approximations. The parameters that we estimate thus do not have a structural interpretation and instead summarize how the function arguments affect firm policies. To be explicit, we substitute into the maximization problem using the following approximations:

$$b(\cdot) \approx \beta_1^A \Delta c_{it} + \beta_2^A N_{it} + \beta_3^A a_{it} + \beta_4^A \bar{c}_{it} + \beta_5^A c_{it}^0 + w'_{it} \alpha^A + \phi_t^A \quad (13)$$

$$-\pi(c_{it}^0, \cdot) \approx \beta_1^S c_{it}^0 + \beta_2^S N_{it} + \beta_3^S a_{it} + \beta_4^S \bar{c}_{it} + w'_{it} \alpha^S + \phi_t^S \quad (14)$$

We specify ϕ_t^A and ϕ_t^S alternately using linear time trends, a flexible polynomial in time, and year fixed effects, which accounts for learning-by-doing and other time-related changes. Most empirical applications of Bajari, Benkard and Levin (2007) use higher-order approximations when estimating policy functions because this helps recover the underlying dynamic structural parameters in the second step. The simpler parameterization is more appropriate here given our empirical objective of understanding the determinants of firm policies.

4.2 Estimation

There is a potentially confounding correlation between the number of firms and the stochastic shocks, which summarize the net effect of unobserved demand and cost factors. To address the issue, we employ the two stage conditional maximum likelihood estimator developed by Rivers and Vuong (1988). The estimator requires a reduced-form equation that governs the evolution of the endogenous variable. We assume that N_{it} evolves according to:

$$N_{it} = z_{it} \gamma_1 + \Delta c_{it} \gamma_2 + a_{it} \gamma_3 + \bar{c}_{it} \gamma_4 + c_{it}^0 \gamma_5 + w'_{it} \gamma_6 + \phi_t^N + v_{it} \quad (15)$$

where z_{it} is an instrument that is excluded from the producers' maximization problem, ϕ_t^N is specified the same way as ϕ_t^A and ϕ_t^S , and v_{it} is a reduced-form error term. For notational convenience, we collect the exogenous variables in the vector X_{it} . We assume that $(X_{it}, u_{it}^A, u_{it}^S, v_{it})$ is i.i.d. Further, let $(u_{it}^A, u_{it}^S, v_{it})$ have a mean-zero joint normal distribution, conditional on X_{it} , with the finite positive definite covariance matrix:

$$\Omega \equiv \begin{bmatrix} \sigma_{uu}^A & \sigma_{uu}^{AS} & \sigma_{vu}^A \\ \sigma_{uu}^{AS} & \sigma_{uu}^S & \sigma_{vu}^S \\ \sigma_{vu}^A & \sigma_{vu}^S & \sigma_{vv} \end{bmatrix} \quad (16)$$

Endogeneity is present if the reduced-form error term is correlated with the stochastic shocks (specifically, if $\sigma_{vu}^A \neq 0$ or $\sigma_{vu}^S \neq 0$). Using the joint normality assumption, the stochastic shocks can be rewritten as $u_{it}^k = v_{it}\lambda^k + \eta_{it}^k$ for $k \in \{A, S\}$, where $\lambda^k = \sigma_{vu}^k/\sigma_{vv}$ and $\eta_{it}^k = u_{it}^k - v_{it}\lambda^k$. If a suitable control function is used as a proxy for the reduced-form error, v_{it} , then the measure of competition is orthogonal to the remaining error terms (Rivers and Vuong (1988)). Estimation proceeds in two stages:

1. OLS estimation of N_{it} on the exogenous regressors. This obtains an estimate of the reduced-form error term that we denote \hat{v}_{it} .
2. Maximum likelihood estimation of the multinomial probit equations using \hat{v}_{it} as a control function. Differences between v_{it} and \hat{v}_{it} are normally distributed and thus compatible with the distributional assumptions of the multinomial probit model.

The second-stage standard errors can be adjusted to account for the presence of the estimation of the control function using a multi-step procedure based on the minimum distance estimator of Amemiya (1978) and Newey (1987). This adjustment has virtually no effect in our application, however, so we report the simpler unadjusted standard errors. It also is possible to cluster standard errors at the kiln-level as an *ad hoc* autocorrelation correction, but this too has little effect on the magnitudes of the standard errors.

4.3 Identification and instrument

The theoretical model indicates that technology adoption is more likely under favorable profit conditions, and greater profit generally supports more competitors. It follows that any correlation between u_{it}^A and v_{it} is likely positive, and this allows the bias to be signed: the basic probit estimator is likely to understate the extent that competition deters technology

adoption. Similar logic applies to kiln shutdown, although our regression results indicate that bias in that equation is empirically less important.

Finding an instrument to correct endogeneity bias is not straightforward. Our setting differs from more standard industrial organization applications that involve demand or supply estimation, and for which cost or demand shocks respectively are valid instruments. Because both demand and cost enter the profit function, neither provides the requisite exogenous variation. We proceed instead under an identifying assumption that competitive conditions exhibit greater autocorrelation than the unobserved profit shocks:

$$\lim_{T \rightarrow \infty} \frac{Cov(u_{it}^A, u_{i,t-T}^A)}{Cov(N_{it}, N_{i,t-T})} = 0 \quad (17)$$

That N_{it} exhibits a high degree of persistence is clear from the data: a regression of N_{it} on $N_{i,t-1}$ and all of the exogenous variables of equation (13) yields a coefficient on the lag of 0.70 (p -value of 0.000). This persistence is due to the longevity of kilns, which are on average 40 years old upon retirement. The degree of persistence in the unobserved profit shock is impossible to evaluate independently, but the frequent changes in observed cost and demand conditions (e.g., see Table 2 and Figure 7) provide some support for this assumption.

Under equation (17), if a lagged version of the competition measure is used as an instrument then bias converges to zero with the length of the lag. We use a lag of 20 years, which exploits the decades of pre-adoption data collected in Chicu (2012). The instrument has power and generates F -statistics between 2,700 and 3,800 in our baseline regressions. Some bias remains if $Cov(u_{it}^A, u_{i,t-20}^A) \neq 0$. This possibility motivates robustness checks in which we use alternative instruments based on respective lags of 15, 10, and 5 years. The results support that the persistence of the error term dies out over longer time horizons.

Other sources of endogeneity seem unlikely. Exogeneity of the demand controls is likely to be reasonable. Technology decisions within a market are not likely to drive demand, because cement represents a small fraction of total construction costs. Endogeneity in fossil fuel prices could arise if increases in fuel demand from cement plants led to price increases in the fuel market. However, any such feedback should be small because cement accounts for a fraction of the fossil fuels used in the United States. Consistent with this argument, bituminous coal prices do not exhibit the same pro-cyclical variation as cement demand. Industry costs (i.e., \bar{c}_{it}) incorporate previous technology decisions and thus could be related to the unobserved profit shocks. However, we obtain similar results if we instrument for industry costs using a 20-year lag on the count of nearby precalciners, and the main results also are robust to the exclusion of industry costs as an independent variable.

5 Variables and summary statistics

5.1 Variables

We calculate the fuel costs of each kiln based on its energy requirements and the price of the primary fuel:

$$\text{Fuel Cost}_{jt} = \text{Primary Fuel Price}_{jt} \times \text{Energy Requirements}_{jt}$$

where the fuel price is in dollars per mBtu and the energy requirements are in mBtu per metric tonne of clinker. We obtain the energy requirements from the PCA labor-energy input surveys. Details on this calculation are provided in Appendix A. The cost savings that would be realized by adopting precalciner technology are the difference between the fuel costs of the kiln and those of the technology frontier, which we define based on the energy requirements of a precalciner kiln using the most affordable fuel. This difference provides the empirical proxy for the Δc term that appears in the theoretical model.

We measure competition based on plant locations and gasoline prices. We first define a distance metric as the multiplicative product of miles and a gasoline price index that equals one in the year 2000. We then calculate the number of competing plants within a distance radius of 400 to obtain an empirical proxy for N_{it} . This radius is motivated by prior findings that 80-90 percent of portland cement is trucked less than 200 miles (Census Bureau (1977); Miller and Osborne (2014)), so that plants separated by a distance of more than 400 are unlikely to compete for customers.²⁰ We exclude plants owned by the same firm from the competition measure, though few such plants exist within the specified radius. We also use the distance radius to calculate the industry costs for each plant (i.e., \bar{c}_{it}), defined as the average costs of all other plants with the radius. The competition and industry cost variables use the location of plants as of the prior year.

Figure 8 provides separate decadal histograms for the count of nearby competitors. Cross-sectional variation is due to plant location dispersion, while inter-temporal variation arises due to gasoline price fluctuations and net plant number decreases over the sample period. We use instruments based on the locations of plants 20 years prior to the observation

²⁰Our treatment of distance reflects the predominant role of trucking in cement distribution. A fraction of cement is shipped to terminals by train (6 percent in 2010) or barge (11 percent in 2010), and only then is trucked to customers. Some plants may therefore be closer than our metric indicates if, for example, both are located on the same river system. Straight-line miles are highly correlated with both driving miles and driving time and, consistent with this, previously published empirical results on the industry are not sensitive to which of these measures is employed (e.g., Miller and Osborne (2014)).

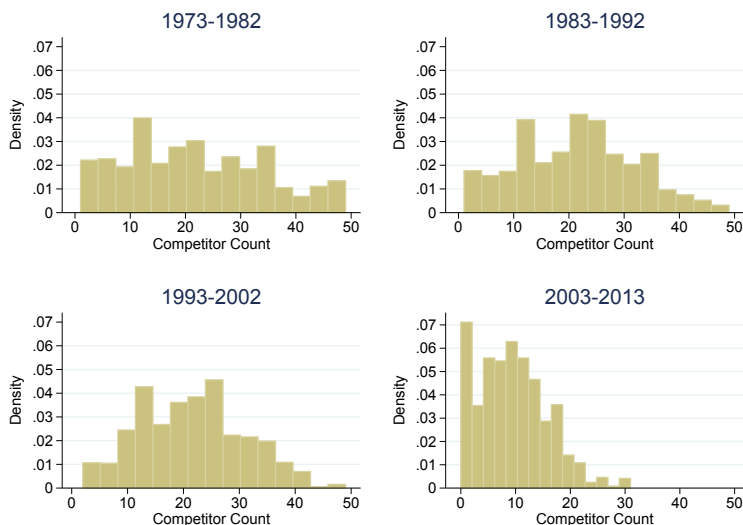


Figure 8: Count of Competitors within a Distance of 400 by Decade

in question. Because gasoline prices are plausibly exogenous, we use the same distance radii to calculate the competition and lagged competition measures. To illustrate, consider a kiln observation in the year 2000, when the gasoline index equals one: instruments are constructed based on the plants in 1980 within 400 miles of the kiln’s location, even though the 1980 gasoline index differs from one. We calculate the instrument in this manner even for kilns that are not present in the data 20 years prior.

Finally, we control for kiln age, kiln capacity, and demand conditions. The two kiln-level controls are straight-forward and obtained from the PIS kiln data. The demand-level control uses county-level data on building permits and construction employment, which explains nearly 90 percent of the variation in USGS-reported state-level consumption. To obtain a single regressor, we first create a county-specific demand variable as a linear combination of building permits and construction employment. The specific formula, which we estimate based on the state-level regressions, is $DEMAND = 0.0154 \times PER + 0.0122 \times EMP$, where PER and EMP are building permits and construction employment, respectively. We then sum the demand among counties within the distance radii from each kiln. As a robustness test, we also constructed variables that capture the distance between plants and the nearest customs district through which foreign imports enter: these controls have little explanatory power, however, and we thus omit them from the specifications shown below.

Table 3: Number of Observations per Kiln

	Count	Mean Obs.	Order Statistics: Observations				
			10%	25%	50%	75%	90%
All Kilns	460	17.81	2	6	12	34	41
Replaced Kilns	144	15.39	2	5	8	28	34
Shut Down Kilns	244	12.82	2	4	10	16	36
Surviving Kilns	72	37.57	37	41	41	41	41

Notes: The table provides the count of unique non-precalciner kilns in the 1973-2013 data, both together and separately for (i) kilns replaced with a precalciner kiln, (ii) kilns closed without replacement, and (iii) kilns in operation as the end of sample period. The table also summarizes the distribution of (annual) observations per kiln.

5.2 Summary statistics

Table 3 describes the sample composition. The data include observations on 460 distinct non-precalciner kilns: 144 are replaced with precalciner technology, 244 are closed without replacement, and 72 survive to the end of the sample. A kiln that is replaced or shut down exits the sample but continues to affect the *Competition* variable for the kilns that remain in the sample. The median kiln is observed for 12 years. At the median, kilns that are replaced with precalciner technology are observed for eight years, kilns that are shut down are observed for ten years, and kilns that survive to the end of the sample are observed for 41 years. There is some variation in the number of observations for surviving kilns due to (infrequent) greenfield entry. There are 8,192 kiln-year observations in the total sample.

Table 4 provides summary statistics for the dependent variables (indicators for adoption and shutdown) and the explanatory variables. Precalciner adoption and kiln shutdown are rare events: indicator means imply an empirical probability of 1.8 percent and 3.0 percent, respectively. The bivariate correlation coefficients show that there are limits to what can be identified given the available empirical variation. Three restrictions on equations (13) and (14) facilitate estimation:

1. We impose that $\beta_5^A = 0$ because the effects of fuel costs and cost savings in the upgrade equation are not separately identifiable, due to the high degree of correlation between the two variables ($\rho = 0.89$). Our analysis is thus focused on the effect of cost savings, which the theoretical model suggests is more important.
2. We impose that $\beta_4^S = 0$ because fuel costs and industry costs are highly correlated ($\rho = 0.86$). We identify only the net effect on the shut down decision.

Table 4: Summary Statistics

Variable	Mean	St. Dev	Correlation Coefficients									
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
(1) Adoption	0.018	0.13										
(2) Shutdown	0.030	0.17	-0.02									
(3) Fuel Cost	c_{it}^0	22.15	9.63	0.07	0.06							
(4) Cost Savings	Δc_{it}	7.78	6.62	0.08	0.05	0.89						
(5) Industry Costs	\bar{c}_{it}	21.68	8.58	0.03	0.06	0.86	0.62					
(6) Competitors	N_{it}	20.56	12.34	-0.00	-0.00	0.03	0.13	0.10				
(7) Construction	a_{it}	12.85	8.85	0.00	-0.04	-0.37	-0.20	-0.38	0.71			
(8) Kiln Age	$w_{1,it}$	30.87	16.12	0.08	0.09	-0.17	-0.09	-0.27	-0.13	0.02		
(9) Kiln Capacity	$w_{2,it}$	0.26	0.18	-0.05	-0.10	-0.20	-0.20	-0.20	-0.01	0.13	-0.39	

Notes: The table provides means, standard deviations, and correlation coefficients for the dependent variables (indicators for adoption and shutdown) and the regressors. The regression sample is comprised of 8,192 kiln-year observations over the period 1973-2013. Capacity is in millions of metric tonnes per year.

3. We impose $\beta_4^A = 0$ in some regressions. If flexible time controls are included (e.g., high-order polynomials or year fixed effects) then the remaining empirical variation is insufficient to identify the effects of industry costs on the upgrade decision.

With these restrictions in hand, there is ample empirical variation to identify the remaining parameters. One way to assess whether collinearity could be problematic is to calculate the variance inflation factors (VIFs) of the regressors. This is done by regressing each regressor k on the other regressors, and calculating $VIF(k) = \frac{1}{1-R^2}$. A rule of thumb is that collinearity is a threat to asymptotic consistency if the VIF exceeds ten (Mela and Kopalle (2002)). In the regressions below, none of the regressors has a VIF that exceeds four.

6 Results

6.1 Baseline regression results

Table 5 presents the baseline probit results. Panel A addresses the likelihood of precalciner adoption and Panel B addresses the likelihood of kiln shutdown. Both are relative to the alternative of maintaining the older kiln. The columns account for changes over time in different ways: column (i) relies exclusively on the regressors; column (ii) adds a linear time

trend ($t = 0, 1, \dots, 40$); column (iii) uses a fifth order polynomial in time; and column (iv) incorporates year fixed effects. The results in column (iv) are generated with two binomial probit regressions due to convergence problems with the multinomial probit.

The results in Panel A show that the likelihood of adoption decreases with the number of nearby competitors. The effect is statistically significant in each regression. To evaluate magnitudes, we calculate that the mean elasticity of the adoption probability with respect to the competition ranges from -1.41 to -2.39 . Thus, interpreted through the lens of the theoretical model, the results indicate that Schumpeterian effects dominate preemption in the data. It is possible that preemption incentives are weak because adoption probabilities are small, or because there are enough plants that the action of any single plant is immaterial for the others. The other parameter estimates are entirely consistent with the comparative statics of the theoretical model: adoption increases with cost savings, construction activity, and industry average costs. The first stage residual has a positive and statistically significant on adoption, which supports our interpretation of the unobserved error term. The mean elasticity with respect to cost savings ranges from 0.58 to 0.71 , and the mean elasticity with respect to nearby construction activity ranges from 1.13 to 1.89 .

Turning to Panel B, the likelihood of kiln shutdown increases in fuel costs and the number of nearby competitors, and decreases with construction activity. Comparing across columns, the precision of the coefficients diminishes with the more flexible controls for time effects. In column (ii), where coefficients remain significant, the mean elasticity of the shutdown probability with respect to fuel costs is 0.41 , the mean elasticity with respect to competitors is 0.74 , and the mean elasticity with respect to construction activity is -0.81 . We suspect that the relatively weaker statistical significance with shutdown arises because shutdown decisions are subject to a greater number of unobservable forces outside of the model. One example would be the exhaustion of an adjacent limestone quarry, which would make continued kiln operations uneconomical. Lastly, the first stage residual does not affect shutdown, so the impact of bias appears to be unimportant empirically.

Among the control variables, kiln age is positively associated both with adoption and shutdown, consistent with common intuition. By contrast, kiln capacity tends to have a negative effect on adoption, though this is statistically significant only if controls for time effects are omitted. At first blush, this result may seem inconsistent with the Schumpeterian hypothesis. However, the regressor does not provide a suitable test because it captures a decision made decades earlier (on average), and because the capacity of the plant changes with precalciner installation. Capacity does correlate negatively with kiln shutdown.

Table 5: Baseline Probit Regression Results

Panel A: Adopt vs. Maintain					
Regressor		(i)	(ii)	(iii)	(iv)
<i>Fuel Costs, Competition, and Demand</i>					
Cost Savings	Δc_{it}	0.051*** (0.007)	0.045*** (0.007)	0.040*** (0.007)	0.032*** (0.006)
Competitors	N_{it}	-0.055*** (0.010)	-0.058*** (0.011)	-0.037*** (0.009)	-0.026*** (0.008)
Construction	a_{it}	0.070*** (0.012)	0.078*** (0.010)	0.049*** (0.011)	0.036*** (0.010)
Industry Costs	\bar{c}_{it}	0.027*** (0.010)	0.029*** (0.010)		
<i>Control Variables</i>					
Kiln Age	$w_{1,it}$	0.018*** (0.003)	0.021*** (0.004)	0.019*** (0.004)	0.014*** (0.003)
Kiln Capacity	$w_{2,it}$	-0.805** (0.392)	-0.675 (0.445)	-0.719 (0.447)	-0.418 (0.341)
First Stage Residual	\hat{v}_{it}	0.075*** (0.014)	0.065*** (0.017)	0.083*** (0.019)	0.050*** (0.014)
<i>Mean Elasticities of Pr(Adoption)</i>					
WRT Cost Savings		0.71	0.63	0.58	0.70
WRT Competitors		-2.25	-2.39	-1.57	-1.41
WRT Construction		1.79	1.89	1.28	1.13
<i>Specification Details</i>					
Time Polynomial		no	1st Order	5th Order	no
Year Fixed Effects		no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regressions in column (iv). The sample is comprised of 8,192 kiln-year observations over 1973-2013. The dependent variable in Panel A is an indicator that equals one if the kiln is replaced with precalciner technology. The dependent variable in Panel B is an indicator that equals one if the kiln is shut down without replacement. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The elasticities are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

Table 5: Baseline Probit Regression Results (continued)

Panel B: Shut Down vs. Maintain					
Regressor		(i)	(ii)	(iii)	(iv)
<i>Fuel Costs, Competition, and Demand</i>					
Fuel Costs	c_{it}^0	0.011** (0.005)	0.011** (0.005)	0.005 (0.007)	-0.002 (0.006)
Competitors	N_{it}	0.017** (0.007)	0.017*** (0.008)	0.021*** (0.008)	0.012* (0.006)
Construction	a_{it}	-0.031*** (0.011)	-0.030*** (0.011)	-0.029*** (0.011)	-0.014 (0.009)
<i>Control Variables</i>					
Kiln Age	$w_{1,it}$	0.016*** (0.003)	0.016*** (0.003)	0.017*** (0.003)	0.012*** (0.002)
Kiln Capacity	$w_{s,it}$	-1.906*** (0.415)	-1.786*** (0.457)	-1.919*** (0.472)	-1.543*** (0.361)
First Stage Residual	\hat{v}_{it}	-0.012 (0.011)	-0.014 (0.012)	-0.014 (0.013)	-0.003 (0.010)
<i>Mean Elasticities of Pr(Shut Down)</i>					
WRT Fuel Costs		0.42	0.41	0.18	0.13
WRT Competitors		0.74	0.72	0.87	0.60
WRT Construction		-0.83	-0.81	-0.79	-0.44
<i>Specification Details</i>					
Time Polynomial		no	1st Order	5th Order	no
Year Fixed Effects		no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regressions in column (iv). The sample is comprised of 8,192 kiln-year observations over 1973-2013. The dependent variable in Panel A is an indicator that equals one if the kiln is replaced with precalciner technology. The dependent variable in Panel B is an indicator that equals one if the kiln is shut down without replacement. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The elasticities are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

6.2 Induced innovation

The positive effect of cost savings in the adoption equation provides support for the proposition that carbon taxes would have induced faster adoption of precalciner technology. The theoretical model suggests that this effect should be particularly true for firms without many competitors. We test this prediction by incorporating interactions. Table 6 summarizes the results for the adoption decision. The interaction of cost savings and competitors is negative and statistically significant in columns (i)-(iii), consistent with the theory. The mean elasticities are similar in magnitude to those of the baseline specifications. The coefficients in the shutdown equation take the expected signs (Appendix Table C.2). They are not statistically significant independently, though some joint significance exists.

Figure 9 plots how a one standard deviation increases in cost savings and fuel costs affect adoption and shutdown, respectively. Panel A shows that greater cost savings increase the probability of adoption only if competition is not too great. The magnitude of the effect for a monopolist is nearly five percentage points: a large percentage given the unconditional probability of adoption is just 1.8 percent. The magnitude of the effect for a firm facing 30 nearby competitors (roughly the 90th percentile) is much smaller. Panel B indicates that shutdown in response to greater fuel costs tends to happen for kilns with many nearby competitors. The magnitudes are large relative to the unconditional shutdown probability of 3.0 percent. Considered together, the analysis indicates that increases in fossil fuel prices are associated with both more adoption and more exit, and that the amount of competition determines which effect dominates.

6.3 Preemption

The baseline results indicate that plants with more nearby competitors are less likely to adopt precalciner technology. For completeness, we provide three additional empirical checks for preemption in this section. First, if preemption is important then logic suggests that the presence of nearby precalciners should discourage adoption and/or encourage shutdown. We therefore add the number of precalciner competitors within a radius of 400 to the baseline specifications. Two first stages are required because the number of precalciner competitors is endogenous. We use 20-year lags on the number of competitors and the number of precalciner competitors as excluded instruments (both of which have considerable power). Table 7 provides the results for adoption in column (i) and shutdown in column (iv). The number of competitors retains its negative effect on precalciner adoption and its positive effect on shutdown. The coefficients on precalciner competitors take the expected signs but are quite

Table 6: Probit Regression Results with Interaction

	Adopt vs. Maintain			
	(i)	(ii)	(iii)	(iv)
<i>Fuel Cost Savings, Competition, and Demand</i>				
Cost Savings	0.066*** (0.016)	0.069*** (0.017)	0.066*** (0.014)	0.046*** (0.011)
Competitors	-0.040*** (0.012)	-0.039*** (0.014)	-0.014 (0.013)	-0.013 (0.011)
Cost Savings × Competitors	-0.0011* (0.0006)	-0.0015** (0.0007)	-0.0017** (0.0007)	-0.0009 (0.0005)
Construction	0.063*** (0.012)	0.066*** (0.013)	0.038*** (0.012)	0.031*** (0.010)
Industry Costs	0.027*** (0.009)	0.030*** (0.010)		
<i>Derived Statistics: Mean Elasticities of Pr(Adoption)</i>				
WRT Cost Savings	0.59	0.54	0.47	0.55
WRT Competitors	-1.98	-2.03	-1.11	-1.18
WRT Construction	1.61	1.67	0.98	0.99
<i>Specification Details</i>				
Control Variables	yes	yes	yes	yes
Time Polynomial	no	1st Order	5th Order	no
Year Fixed Effects	no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regression in column (iv). The sample is composed of 8,192 kiln-year observations over 1973-2013. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The control variables include kiln age, kiln capacity, the first stage residual, and the first stage residual interacted with cost savings. The elasticities of the estimated adoption probability with respect to *Cost Savings*, *Competitors*, and *Construction* are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

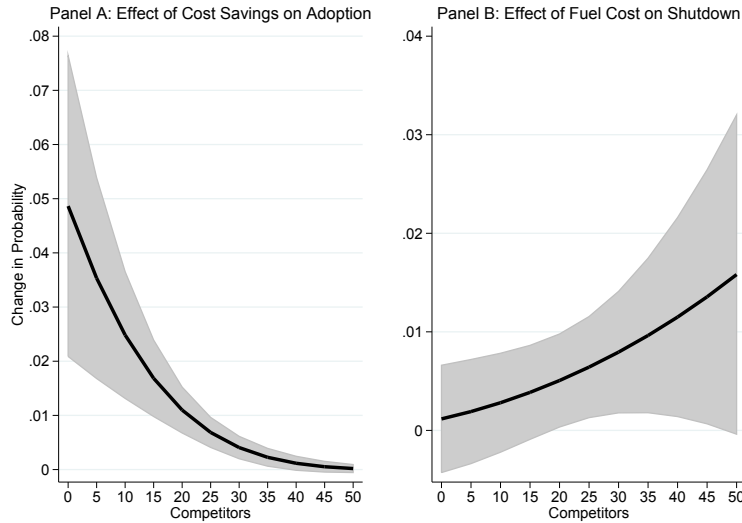


Figure 9: Changes in the Probabilities of Adoption and Shutdown

Notes: Panel A plots changes in precalciner adoption probabilities due to a one standard deviation increase in cost savings (which is 6.62). Panel B plots changes in kiln shutdown probabilities due to a one standard deviation increase in fuel costs (which is 9.63). Both effects are evaluated over the number of nearby competitors. Results are based on specification (ii) in Table 6. Other explanatory variables are held at their respective means. The shaded regions provide 95% confidence intervals.

imprecisely estimated. Thus, the regression does not provide compelling empirical support for the hypothesis that early adoption preempts competitors or encourages shutdown.

Second, the theoretical model indicates that if preemption incentives are strong enough then an inverted-U relationship between competition and adoption arises in the early periods of industry adoption. If the inverted-U is present, however, it is a short-term phenomenon because the theoretical model is unambiguous that competition weakly decreases long run adoption. We test for an inverted-U relationship by allowing the benefits of adoption to have a quadratic relationship with the competition measure. In estimation, we also incorporate a quadratic in the estimated first-stage residual. The results are provided in columns (ii) and (v). The quadratic term is *positive* in the adoption equation, exactly the opposite of what would be arise with an inverted-U relationship. We emphasize that the net effect of the competition regressors is negative in the range of the data, so the results should not be misinterpreted as implying that a sufficiently high degree of competition increases adoption.

A third approach to testing for preemption is to see whether adoption is most likely for moderate levels of demand (e.g., Dafny (2005) Ellison and Ellison (2011)). The logic is that competitors definitely would not adopt with low enough demand and definitely would adopt with high enough demand, which isolates intermediate ranges of demand as candidates for

Table 7: Additional Checks for Preemption

	Adopt vs. Maintain			Shut Down vs. Maintain		
	(i)	(ii)	(iii)	(iv)	(v)	(v)
Cost Savings	0.041*** (0.007)	0.041*** (0.007)	0.041*** (0.007)			
Fuel Costs				0.004 (0.007)	0.005 (0.007)	0.005 (0.007)
Competitors	-0.039*** (0.011)	-0.073*** (0.014)	-0.035*** (0.010)	0.021** (0.009)	0.029** (0.013)	0.011 0.008
Precalciner Competitors	-0.012 (0.059)			0.014 (0.060)		
Competitors ²		0.0007*** (0.0002)			-0.0002 (0.0002)	
Construction	0.052*** (0.013)	0.046*** (0.011)	0.035 (0.024)	-0.029*** (0.012)	0.028** (0.011)	0.051** (0.024)
Construction ²			0.0003 (0.0005)			-0.002*** (0.001)

Notes: Results are from multinomial probit regressions. The data include 8,192 kiln-year observations over 1973-2013. The dependent variable in the left columns is an indicator that equals one if the kiln is replaced with precalciner technology, and the dependent variable in the right columns is an indicator that equals one if the kiln is shut down without being replaced. All regressions incorporate control variables and a fifth order polynomial in time. The control variables are kiln age, kiln capacity, and the first stage residual(s). In columns (i) and (iv), there are two first stage regressions, for the number of competitors and the number of precalciner competitors, respectively. The excluded instruments are 20-year lags on the competition variables. In the other columns there is a single first stage regression and the excluded instrument is a 20-year lag on competition. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

supporting preemption. We implement by adding a quadratic term in *Construction*. The results are shown in columns (iii) and (vi). The coefficient on the quadratic term is positive in the adoption equation. This result does not support preemption, but the standard error is large enough that the opposite relationship cannot be ruled out.

6.4 Robustness analysis

The results developed above are robust to alternative choices related to the distribution of the structural error terms, variable definitions, and the relevant sample period. Regressions that use the binomial probit, binomial logit, and multinomial logit models return basically identical results. The linear probability model also returns similar effects in terms of both magnitude and statistical significance. We also have used a “competing risks” semiparametric hazard rate model (Fine and Gray (1999)), in which shutdown is incorporated as an exogenous event rather than as an endogenous decision driven by particular economic circumstances. Results are consistent with the binomial probit. The estimates appear to be driven by the empirical variation in the data rather than particular distributional assumptions.

Table 8 evaluates robustness with respect to variable definitions and sample periods. We use binomial probit regressions to estimate the adoption equation. Column (i) adds two alternative cost savings measures, based on fossil fuel prices five years ahead and behind the year of the observation. These alternative measures do not predict adoption. Column (ii) modifies the number of competitors based on a tighter distance radius of 200. This alternative competition measure does not affect the results much. Column (iii) uses both the baseline radius (400) and the alternative radius (200), and both variables are found negative and statistically significant. The total effect of a competitor within a radius of 200 is $-0.015 - 0.041 = -0.056$. Closer competitors matter more, which is consistent with the role of transportation costs. Columns (iv) and (v) use subsamples that respectively span 1973-1990 and 1991-2013. The results do not differ substantially.

Table 9 explores the IV strategy. Again we use binomial probit regressions to estimate the adoption equation. Column (i) excludes the first stage residual. The number of nearby competitors still has a negative effect but the magnitude is reduced. The direction of this change is consistent with expectations given the source of bias. Columns (ii)-(iv) respectively use 5-year, 10-year, and 15-year lags on competition as the excluded instrument, instead of the 20-year lags used in the baseline specifications. The magnitude of the estimated effect of nearby competitors is greater if a longer lag is used as an instrument. The mean elasticity of adoption with respect to competition is -1.37 with the 15-year lag, which is not statistically

Table 8: Probit Regressions with Alternative Regressors and Samples

	(i)	(ii)	(iii)	(iv)	(v)
Cost Savings	0.038*** (0.008)	0.015** (0.006)	0.032*** (0.005)	0.030*** (0.005)	0.040** (0.017)
Cost Savings ($t + 5$)	-0.008 (0.005)				
Cost Savings ($t - 5$)	0.000 (0.010)				
Competitors ($d < 400$)	-0.027*** (0.008)		-0.015** (0.007)	-0.028** (0.012)	-0.030** (0.015)
Competitors ($d < 200$)		-0.071*** (0.014)	-0.041*** (0.013)		
Construction	0.030*** (0.011)	0.085*** (0.017)	0.042*** (0.009)	0.041** (0.018)	0.032** (0.013)

Notes: Results are from binomial probit regressions. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The data in columns (i)-(iii) include 8,192 kiln-year observations over 1973-2013. The data in column (iv) include 5,149 kiln-year observations over 1973-1990, and the data in column (v) include 3,043 kiln-year observations over 1991-2010. All regressions incorporate control variables and a fifth order polynomial in time. The control variables are kiln age, kiln capacity, and the first stage residual. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

Table 9: Probit Regression Results with Alternative IVs

	(i)	(ii)	(iii)	(iv)
Cost Savings	0.027*** (0.005)	0.028*** (0.005)	0.028*** (0.005)	0.029*** (0.004)
Competitors	-0.013*** (0.005)	-0.017*** (0.006)	-0.021*** (0.006)	-0.025*** (0.006)
Construction	0.024*** (0.007)	0.029*** (0.008)	0.032*** (0.008)	0.034*** (0.008)
Residual		0.063*** (0.025)	0.059*** (0.018)	0.053*** (0.014)
IV Lag Structure	No IV	5 Year	10 Year	15 Year
<i>Derived Statistics: Mean Elasticities of Pr(Adoption)</i>				
WRT Cost Savings	0.52	0.54	0.55	0.56
WRT Competitors	-0.69	-0.92	-1.15	-1.37
WRT Construction	0.80	0.98	1.07	1.16

Notes: Results are from binomial probit regressions. The dependent variable is an indicator that equals one if the kiln is replaced with precalciner technology. The data include 8,192 kiln-year observations over 1973-2013. All regressions incorporate control variables and a fifth order polynomial in time. The control variables are kiln age, kiln capacity, and the first stage residual. The excluded instrument in the first stage is a lag on the number of nearby competitors, as described within the columns. The elasticities of the estimated adoption probability with respect to *Cost Savings*, *Competitors*, and *Construction* are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.

different from the -1.57 mean elasticity obtained with the comparable baseline specification. Thus, while the results suggest even longer lags might be required to fully eliminate bias, the effects of persistence in the error term appear to mostly die out within 15-20 years.

7 Conclusion

Our research explores precalciner adoption in the portland cement industry. The precalciner kiln reasonably can be characterized as belonging to a broad class of non-drastic, non-exclusive, and non-divisible cost-reducing technologies. A theoretical analysis of this setting shows that adoption incentives exhibit a Schumpeterian flair: competition can deprive firms of the scale necessary to recoup fixed adoption costs. The empirical analysis of firm policy functions comports with comparative statics of the theoretical model and indicates that plants with fewer nearby competitions are more likely to adopt precalciner technology.

The fundamental argument of Schumpeter (1942) is that market power in capitalist economies creates conditions under which “the perennial gale of creative destruction” facilitates innovation and growth. Thus, it may seem ironic that support for the Schumpeterian hypothesis is found in the staid industry of portland cement, which has used broadly similar production technologies for more than a century. On the other hand, three generations of Schumpeter’s ancestors ran a successful textiles company, at one point installing the first steam engine in Triesch in modern-day Czech Republic; at least one biographer speculates this family history contributed to Schumpeter’s interests in entrepreneurship and innovation McCraw (2007, p. 11). Furthermore, it is precisely the stability of the portland cement industry that makes it amenable to careful empirical analysis. In this sense, our research is similar to other studies that have found support for Schumpeterian hypotheses in the diffusion of process innovations (e.g., Hannan and McDowell (1984); Rose and Joskow (1990); Karshenas and Stoneman (1993); Akhavein, Frame and White (2005); Schmidt-Dengler (2006)).

We conclude with a brief discussion of possible extensions. Our current research focuses on how and why competition affects firm policies regarding technology adoption. We have not (yet) examined welfare effects. It also is possible to incorporate entry and exit into the model. Our early efforts on this front indicate that preemption effects become somewhat stronger with exit, but that otherwise the comparative statics are unchanged. We have opted to present the simpler model here. The main advantage to the richer model is that it is more suitable for structural analyses of industry adjustment paths. Such an undertaking would require simplifying restrictions on the state space, but if these can be made without too much loss of realism then useful results could be obtained on a number of subjects, including the welfare consequences of market-based regulation.

References

- Abernathy, William J. and James M. Utterback**, “Patterns of Industrial Innovation,” *Technology Review*, June/July 1978, pp. 41–47.
- Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr**, “Transition to Clean Technology,” *Journal of Political Economy*, 2016, *124* (1), 52–104.
- Aghion, Philippe and Jean Tirole**, “The Management of Innovation,” *Quarterly Journal of Economics*, 1994, *109*, 1185–1209.
- and **Rachel Griffith**, *Competition and Growth: Reconciling Theory and Evidence*, The MIT Press, 2005.
- , **Antoine Dechezlepretre, David Hemous, Ralf Martin, and John Van Reenen**, “Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry,” 2012. NBER Working Paper 18596.
- , **Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt**, “Competition and Innovation: An Inverted U-Relationship,” *Quarterly Journal of Economics*, 2005, *120*, 701728.
- Akhavein, Jalal, W. Scott Frame, and Lawrence J. White**, “The Diffusion of Financial Innovations: An Examination of the Adoption of Small Business Credit Scoring in Large Banking Organizations,” *Journal of Business*, 2005, *78* (2), 577–596.
- Amemiya, Takesh**, “The Estimation of a Simultaneous Equation Generalized Probit Model,” *Econometrica*, 1978, *46*, 1193–1205.
- Arrow, Kenneth J.**, “Economic Welfare and the Allocation of Resources for Invention,” 1962. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. J. Neyman, 507-532. Berkeley: University of California Press.
- Bajari, P., L. Benkard, and J. Levin**, “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 2007, pp. 1331–1370.
- Baldwin, William L. and John T. Scott**, “Market Structure and Technological Change,” 1987. In Lesourne and Sonnenschein (Eds.), *Fundamentals of Pure and Applied Economics*. Chur, Switzerland, and London: Harwook Academic Publishers.

- Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer**, “Are Consumers Myopic? Evidence from New and Used Car Purchases,” *American Economic Review*, 2013, 103 (1), 220–256.
- Carbon War Room**, “Gigatron Analysis of the Cement Industry,” 2011.
- Chicu, Mark**, “Dynamic Investment and Deterrence in the U.S. Cement Industry,” 2012. Job Market Paper, Northwestern University.
- Cohen, Wesley M.**, “Empirical Studies of Innovative Activity,” 1995. In Stoneman, P. (Ed.), *Handbook of the Economics of Innovation and Technical Change*. Oxford: Basil Blackwell.
- , “Fifty Years of Empirical Studies of Innovative Activity and Performance,” 2010. In Arrow, K.J. and Intriligator, M.D. (eds), *Handbook of the Economics of Innovation*, Volume 01. Elsevier.
- and **Richard C. Levin**, “Empirical Studies of Innovation and Market Structure,” 1989. In Schmalensee, R. and Willig, R. (Eds.), *Handbook of Industrial Organization*. Amsterdam: North-Holland.
- Dafny, Leemore**, “Games that Hospitals Play: Entry Deterrence in Hospital Procedure Markets,” *Journal of Economics & Management Strategy*, 2005, 14 (3), 513–542.
- Dasgupta, Partha and Joseph E. Stiglitz**, “Industrial Structure and the Nature of Innovative Activity,” *The Economic Journal*, 1980, 90 (358), 266–293.
- Doraszelski, Ulrich and Mark Satterthwaite**, “Computable Markov-Perfect Industry Dynamics,” *RAND Journal of Economics*, 2010, 41, 215–243.
- Ellison, Glenn and Sara Fisher Ellison**, “Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbent Prior to Patent Expiration,” *American Economic Journal: Microeconomics*, 2011, 3 (1), 1–36.
- Ericson, Richard and Ariel Pakes**, “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, 1995, 62 (1), 53–82.
- Fang, Limin**, “Measuring Preemption in Dynamic Oligopoly Games,” 2016.
- Fine, J. P. and R. J. Gray**, “A Proportional Hazards Model for the Subdistribution of a Competing Risk,” *Journal of the American Statistical Association*, 1999, 94, 496–509.

- Fowlie, Meredith, Mar Reguant, and Stephen P. Ryan**, “Market-Based Emissions Regulation and Industry Dynamics,” *Journal of Political Economy*, 2016, *124* (1), 249–302.
- Fudenberg, Drew and Jean Tirole**, “A Theory of Exit in Duopoly,” *Econometrica*, 1986, *54* (4), 943–960.
- Fuentes, Roberto, Ruben Hernandez-Murillo, and Gerard Llobet**, “Strategic Online-Banking Adoption,” *Journal of Banking and Finance*, 2010, *34*, 1650–1663.
- Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker**, “Energy Prices, Pass-Through, and Incidence in U.S. Manufacturing,” 2016.
- Genesove, David**, “The Adoption of Offset Presses in the Daily Newspaper Industry in the United States,” 1999. NBER Working Paper No. 7076.
- Ghemaway, Pankaj and Barry Nalebuff**, “Exit,” *RAND Journal of Economics*, 1985, *16* (2), 184–194.
- Gil, Ricard, Jean-François Houde, and Yuya Takahashi**, “Preemptive Entry and Technology Diffusion in the Market for Drive-In Theaters,” 2015. Working Paper.
- Gilbert, Richard**, “Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate?,” 2006. In Jaffe, A.B., Lerner, J. and Stern, S. (Eds.), *Innovation Policy and the Economy*. Boston: The MIT Press, 159-215.
- and **David Newbery**, “Preemptive Patenting and the Persistence of Monopoly,” *American Economic Review*, 1982, *72*, 514–526.
- Goettler, Ronald and Brett Gordon**, “Does AMD Spur Intel to Innovate More?,” *Journal of Political Economy*, 2011, *119* (6), 1141–1200. Mimeo.
- Griliches, Zvi**, “Hybrid Corn: An Exploration in the Economics of Technological Change,” *Econometrica*, 1957, *25* (4), 501–522.
- Hanlon, W. Walker**, “Necessity is the Mother of Invention: Input Supplies and Directed Technical Change,” 2014.
- Hannan, Timothy H. and John M. McDowell**, “The Determinants of Technology Adoption: The Case of the Banking Firm,” *RAND Journal of Economics*, 1984, *15* (3), 328–335.

- Igami, Mitsuru**, “Estimating the Innovator’s Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981-1998,” *Journal of Political Economy*, 2015. Forthcoming.
- and **Kosuke Uetake**, “Mergers, Innovation, and Entry-Exit Dynamics: The Consolidation of the Hard Disk Drive Industry, 1996-2015,” 2016.
- Iskhakov, Fedor, John Rust, and Bertel Schjerning**, “The Dynamics of Bertrand Price Competition with Cost-Reducing Investments,” 2015. Working Paper.
- Kamien, Morton I. and Nancy L. Schwartz**, *Market Structure and Innovation*, Cambridge: Cambridge University Press, 1982.
- Karshenas, Massoud and Paul L. Stoneman**, “Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model,” *RAND Journal of Economics*, 1993, *24* (4), 203–528.
- Langer, Ashley and Nathan H. Miller**, “Automakers’ Short-Run Responses to Changing Gasoline Prices,” *Review of Economics and Statistics*, 2013, *95* (4), 1198–1211.
- Linn, Joshua**, “Energy Prices and the Adoption of Energy-Saving Technology,” *The Economic Journal*, 2008, *118*, 1986–2012.
- McCraw, Thomas K.**, *Prophet of Innovation: Joseph Schumpeter and Creative Destruction*, Cambridge, MA: The Belknap Press of Harvard University Press, 2007.
- Mela, Carl F. and Praveen K. Kopalle**, “The Impact of Collinearity on Regression Analysis: The Asymmetric Effect of Negative and Positive Correlations,” *Applied Economics*, 2002, *34*, 667–677.
- Miller, Nathan H. and Matthew Osborne**, “Spatial Differentiation and Price Discrimination in the Cement Industry: Evidence from a Structural Model,” *RAND Journal of Economics*, 2014, *45* (2), 221–247.
- , — , and **Gloria Sheu**, “Pass-Through in a Concentrated Industry: Empirical Evidence and Regulatory Implications,” 2015.
- Newell, Richard G., Adam B. Jaffe, and Robert N. Stavins**, “The Induced Innovation Hypothesis and Energy-Saving Technological Change,” *Quarterly Journal of Economics*, 1999, *114* (3), 941–975.

- Newey, Whitney**, “Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables,” *Journal of Econometrics*, 1987, *36*, 231–250.
- Oster, Sharon**, “The Diffusion of Innovation among Steel Firms: The Basic Oxygen Furnace,” *The Bell Journal of Economics*, 1982, *13* (1), 45–56.
- Popp, David**, “Induced Innovation and Energy Prices,” *American Economic Review*, 2002, *92* (1), 160–181.
- Rivers, Douglas and Quang H. Vuong**, “Limited Information Estimators and Exogeneity TTest for Simultaneous Probit Models,” *Journal of Econometrics*, 1988, *39*, 347–366.
- Rose, Nancy and Paul L. Joskow**, “The Diffusion of New Technologies: Evidence from the Electric Utility Industry,” *RAND Journal of Economics*, 1990, *21* (3), 354–373.
- Ryan, Stephen**, “The Costs of Environmental Regulation in a Concentrated Industry,” *Econometrica*, 2012, *80* (3), 1019–1062.
- Schmidt-Dengler, Philip**, “The Timing of New Technology Adoption: The Case of MRI,” 2006. Working Paper.
- Schumpeter, Joseph A.**, *The Theory of Economic Development: An Inquiry Into Profits, Capital Credit, Interest, and the Business Cycle*, Cambridge, MA: Harvard University Press, 1934.
- , *Capitalism, Socialism, and Democracy*, New York, NY: Harper, 1942.
- Shapiro, Carl**, “Theories of Oligopoly Behavior,” 1989. In *Handbook of Industrial Organization, volume 1*, Richard Schmalensee and Robert Willing (eds). Elsevier.
- Takahashi, Yuya**, “Estimating a War of Attrition: The Case of the US Movie Theater Industry,” *American Economic Review*, 2015, *105* (7), 2204–2241.
- Van Oss, Hendrik G.**, “Background Facts and Issues Concerning Cement and Cement Data,” 2005. Open-File Report 2005-1152, U.S. Department of the Interior, U.S. Geological Survey.
- Vogt, William B.**, “Detecting Strategic Behavior in Technology Adoption: The Example of Magnetic Resonance Imaging,” 2000. Unpublished working paper.

Appendices

A Measuring Fuel Costs

We calculate the energy requirements of production based on the labor-energy input surveys of the PCA. There is no discernible change in the requirements over 1990-2010, conditional on the kiln type. We calculate the average mBtu per metric tonne of clinker required in 1990, 2000, and 2010, and apply these averages over 1990-2013. These are 3.94, 4.11, 5.28, and 6.07 mBtu per metric tonne of clinker for precalciner kilns, preheater kilns, long dry kilns, and wet kilns, respectively. A recent USGS survey accords with our calculations (Van Oss (2005)). Technological improvements are evident over 1973-1990 within kiln type: in 1974, the energy requirements were 6.50 mBtu per metric tonne of clinker at dry kilns (a blended average across dry kiln types) and 7.93 mBtu per metric tonne of clinker at wet kilns. We assume that improvements are realized linearly over 1973-1990. We scale down by our calculated energy requirements by five percent to reflect that a small amount of gypsum is ground together with the kiln output (i.e., clinker) to form cement.

Plants sometimes list multiple primary fuels in the PIS. In those instances, we calculate fuel costs with the coal price if coal is among the primary fuels; otherwise, we use natural gas prices if natural gas is among the multiple fuels. We use oil prices only if oil is the only fossil fuel listed. In the 1980s, petroleum coke supplements or replaces coal at many kilns. The price of coal and petroleum coke are highly correlated, and we simply use the coal price for those observations. Figure A.1 plots fossil fuel prices and usage over the sample period. In the mid-1970s, coal and natural gas were the most popular fuel choices, while only a small subset of plants used oil. Coal quickly came to dominate the industry due to a change in relative prices, and fuel costs thereafter track the coal price.

Our methodology does not incorporate secondary fuels, the most popular of which are waste fuels such as solvents and used tires. The labor-energy input surveys of the PCA indicate that waste fuels account for around 25% of the energy used in wet kilns and 5% of the energy used in dry kilns. We do not have data on the prices of waste fuels but understand them to be lower on a per-mBtu basis than those of fossil fuels. Accordingly, we construct an alternative fuel cost measure in which we scale down the fossil fuel requirements of wet and dry kilns in accordance with the survey data. Whether this adjustment better reflects the fuel costs of marginal output depends in part on (i) the relative prices of waste and fossil fuels and (ii) whether the average fuel mix reported in the survey data reflect the marginal fuel mix. On the latter point, if marginal clinker output is fired with fossil fuels then our

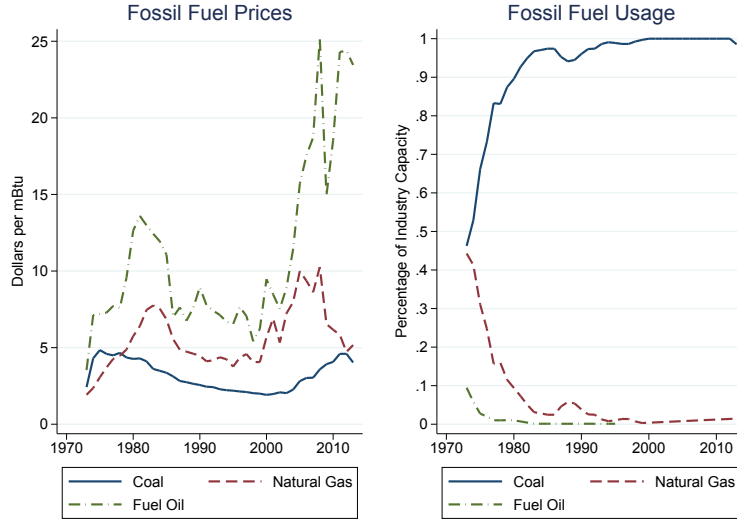


Figure A.1: Fossil Fuel Prices and Usage 1973-2013

baseline measurement should reflect marginal fuel costs more closely than the alternative measurement. Regardless, our regression results are not very sensitive to the adjustment.

B Theory

B.1 Intermediate calculations

This section provides the intermediate steps needed to derive the value functions of equation (3) and (4), the benefit function of equation 5, and the condition under which competition reduces the benefit of adoption (i.e., equation (9)). First, it is simple to solve explicitly for the future value of adoption, $\mathbf{V}_1(N)$ because once the adoption decision is irreversible. In particular, the value function vector can be written as

$$\mathbf{V}_1(N) = \mathbf{\Pi}_0 \left(\boldsymbol{\pi}_1 + \delta \left(\sum_{t=0}^{\infty} \delta^t \mathbf{\Pi}_1^t \boldsymbol{\pi}_1 \right) \right) \quad (\text{B.1})$$

$$= \mathbf{\Pi}_0 \left(\mathbf{I} + \delta(\mathbf{I} - \delta \mathbf{\Pi}_1)^{-1} \right) \boldsymbol{\pi}_1, \quad (\text{B.2})$$

and this is equation (3). We have re-expressed profit such that $\boldsymbol{\pi}_1 = \boldsymbol{\pi}(c_1, N)$ for notional brevity, and we also will use $\boldsymbol{\pi}_0 = \boldsymbol{\pi}(c_0, N)$ in this appendix. Second, although it is slightly more complicated to solve for $\mathbf{V}_0(N)$, we can apply a similar idea to the one above. The

$(L + 1)^{th}$ element of $\mathbf{V}_0(N)$ is given by

$$V_0(L; N) = F(k^*(L))v_1(L; N) + (1 - F(k^*(L)))v_0(L; N) \quad (\text{B.3})$$

Plugging in for $v_1(\cdot)$ and $v_0(\cdot)$ based on equation (2) yields

$$\begin{aligned} V_0(L; N) &= \mathbf{P}_0(L)' \left(F(k^*(L))(\boldsymbol{\pi}_1 + \delta\mathbf{V}_1(N)) + (1 - F(k^*(L)))(\boldsymbol{\pi}_0 + \delta\mathbf{V}_0(N)) \right) \\ &\quad - F(k^*(L))E(k|k < k^*(L)) \end{aligned}$$

The term in the second line of this equation is the conditional expected cost of adoption. Now plugging in for $\mathbf{V}_1(N)$ based on equation (B.2) yields:

$$\begin{aligned} V_0(L; N) &= \mathbf{P}_0(L)' \left(F(k^*(L)) (\mathbf{I} + \delta\boldsymbol{\Pi}_0 (\mathbf{I} + \delta(\mathbf{I} - \delta\boldsymbol{\Pi}_1)^{-1})) \boldsymbol{\pi}_1 \right) \\ &\quad + \mathbf{P}_0(L)' \left((1 - F(k^*(L)))(\boldsymbol{\pi}_0 + \delta\mathbf{V}_0(N)) \right) \\ &\quad - F(k^*(L))E(k|k < k^*(L)) \end{aligned}$$

Stacking the $V_0(L; N)$ elements for $L = 0, \dots, N - 1$ yields an equation that defines $\mathbf{V}_0(N)$:

$$\mathbf{V}_0(N) = \boldsymbol{\Pi}_0 (\mathbf{D} (\mathbf{I} + \delta\boldsymbol{\Pi}_0 (\mathbf{I} + \delta(\mathbf{I} - \delta\boldsymbol{\Pi}_1)^{-1})) \boldsymbol{\pi}_1 + (\mathbf{I} - \mathbf{D})\boldsymbol{\pi}_0 + \delta(\mathbf{I} - \mathbf{D})\mathbf{V}_0) + \mathbf{D}\boldsymbol{\kappa}$$

where \mathbf{D} is a diagonal matrix with $F(k^*(L))$ on each $(L + 1)^{st}$ diagonal element, and $\boldsymbol{\kappa}$ is a vector with $E(k|k < k^*(L))$ for each $(L + 1)^{st}$ element. Solving for \mathbf{V}_0 obtains

$$\begin{aligned} \mathbf{V}_0(N) &= (\mathbf{I} - \delta\boldsymbol{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1}\boldsymbol{\Pi}_0\mathbf{D} (\mathbf{I} + \delta\boldsymbol{\Pi}_0 (\mathbf{I} + \delta(\mathbf{I} - \delta\boldsymbol{\Pi}_1)^{-1})) \boldsymbol{\pi}_1 \\ &\quad + (\mathbf{I} - \delta\boldsymbol{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1}\boldsymbol{\Pi}_0(\mathbf{I} - \mathbf{D})\boldsymbol{\pi}_0 \\ &\quad + (\mathbf{I} - \delta\boldsymbol{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1}\mathbf{D}\boldsymbol{\kappa} \end{aligned} \quad (\text{B.4})$$

The benefit of adoption given state L is defined in the main body of the text as $b(L; N, k) = v_1(L; N, k) - v_0(L; N)$. Stacking across states yields the vector

$$\mathbf{b}(N, k) = \boldsymbol{\Pi}_0 (\boldsymbol{\pi}_1 + \delta\mathbf{V}_1(N) - \boldsymbol{\pi}_0 - \delta\mathbf{V}_0(N)) - k \quad (\text{B.5})$$

The benefit of adoption can be evaluated by plugging in using the value functions defined

in equations (3) and (B.5). This obtains

$$\begin{aligned}
\mathbf{b}(N, k) &= \mathbf{\Pi}_0(\boldsymbol{\pi}_1 - \boldsymbol{\pi}_0) - k \\
&+ \delta \mathbf{\Pi}_0(\mathbf{\Pi}_0(\mathbf{I} + \delta(\mathbf{I} - \delta \mathbf{\Pi}_1)^{-1}))\boldsymbol{\pi}_1 \\
&- \delta \mathbf{\Pi}_0(\mathbf{I} - \delta \mathbf{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1}\boldsymbol{\pi}_0 \\
&- \delta \mathbf{\Pi}_0(\mathbf{I} - \delta \mathbf{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1}[\mathbf{\Pi}_0 \mathbf{D}(\boldsymbol{\pi}_1 - \boldsymbol{\pi}_0) - \mathbf{D}\boldsymbol{\kappa}] \\
&- \delta^2 \mathbf{\Pi}_0((\mathbf{I} - \delta \mathbf{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1} \mathbf{\Pi}_0 \mathbf{D} \mathbf{\Pi}_0(\mathbf{I} + \delta(\mathbf{I} - \delta \mathbf{\Pi}_1)^{-1}))\boldsymbol{\pi}_1
\end{aligned} \tag{B.6}$$

The intermediate matrices that convert this into equation (5) are given by:

$$\begin{aligned}
\mathbf{A}(\mathbf{\Pi}_0, \mathbf{\Pi}_1) &= \mathbf{\Pi}_0(\mathbf{I} + \delta(\mathbf{I} - \delta \mathbf{\Pi}_1)^{-1}) \\
\mathbf{B}(\mathbf{\Pi}_0) &= (\mathbf{I} - \delta \mathbf{\Pi}_0(\mathbf{I} - \mathbf{D}))^{-1}
\end{aligned}$$

and, again, \mathbf{D} is a diagonal matrix with $F(k^*(L))$ on each $(L + 1)^{st}$ diagonal element. If adoption probabilities are small, which is the case in our empirical setting, then $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_1$ are close to the identity matrix and \mathbf{D} is close to the zero matrix. This causes preemption incentives and the option value to be small. Tracing the implications through the intermediate matrices, the benefit can be approximated as

$$\mathbf{b}(N) \approx -\frac{1}{1 - \delta} (\boldsymbol{\pi}_1 - \boldsymbol{\pi}_0) - k, \tag{B.7}$$

We turn now to equation (9). What must be derived is the condition under which $b(L; N) > b(L; N + 1)$, keeping in mind that L enters the $b(\cdot)$ function through the average industry cost \bar{c} . Plugging in based on equations (7) and (8) yields

$$\frac{N + 1}{N + 2} \frac{a - \hat{c} + (N + 1)(\bar{c} - \hat{c})}{N + 2} < \frac{N}{N + 1} \frac{a - \hat{c} + N(\bar{c} - \hat{c})}{N + 1} \tag{B.8}$$

Rearranging the sides of the equation yields

$$\left[\frac{(N + 1)^2}{(N + 2)^2} - \frac{(N)^2}{(N + 1)^2} \right] (\bar{c} - \hat{c}) < \left[\frac{N}{(N + 1)^2} - \frac{(N + 1)}{(N + 2)^2} \right] (a - \hat{c})$$

and a few more lines of algebra obtain the equation shown in the text.

B.2 Proposition 1

Discussion: States in the special class are characterized by a number of adopters L^* such that the $(L^* + 1, L^* + 1)$ element of the industry transition matrix $\tilde{\mathbf{\Pi}}$ equals one, and the $(L^*, L^* + 1)$ element of $\tilde{\mathbf{\Pi}}$ is positive. Suppose that adoption occurs given some industry state and cost draw. Then it follows that “non-strategic” adoption would occur, conceptualizing this as being driven by a cut-off rule that does not account for the possibility that competitors also might adopt. Thus, the absorbing state of interest is defined by some L^* that satisfies

$$(1 - \delta)\underline{k} \leq q^*(c_1, L^*; N)^2 - q^*(c_0, L^* - 1; N)^2 \quad (\text{B.9})$$

and

$$(1 - \delta)\underline{k} > q^*(c_1, L^* + 1; N)^2 - q^*(c_0, L^*; N)^2 \quad (\text{B.10})$$

recalling that $\pi(c_x, L; N) = (q^*(c_x, L; N))^2$ for $x \in \{0, 1\}$. The first inequality states that non-strategic adoption occurs given state $L^* - 1$ and a cost draw of \underline{k} . The latter inequality states that adoption does not occur at state L^* . Its derivation does not require integration over competitor actions because L^* is an absorbing state. We first establish three useful lemmas and state a corollary.

Lemma 1: $q^*(c_1, L; N) = q^*(c_0, L; N) + \Delta c$ for any L .

Proof: The relationship is derived from the expressions for equilibrium quantities provided in Section 2.3. Adding Δc to both sides of equation (7) yields

$$\begin{aligned} q^*(c_0, L; N) + \Delta c &= \frac{a - c_0 + N(\bar{c} - c_0)}{(N + 1)} + \frac{N + 1}{N + 1} \Delta c \\ &= \frac{a - (c_0 - \Delta c) + N(\bar{c} - (c_0 - \Delta c))}{(N + 1)} \\ &= \frac{a - c_1 + N(\bar{c} - c_1)}{(N + 1)} \\ &= q^*(c_1, L; N) \end{aligned}$$

The third line uses the definition $\Delta c = c_0 - c_1$. QED.

Lemma 2: $q^*(c_0, L; N) = \frac{a - c_0 - L\Delta c}{N + 1}$

Proof: The relationship again is derived from the expressions for equilibrium quantities

provided in Section 2.3. From equation (7),

$$\begin{aligned}
q^*(c_0, L; N) &= \frac{a - c_0 + N(\bar{c} - c_0)}{(N + 1)} \\
&= \frac{a - (N + 1)c_0 + N\bar{c}}{(N + 1)} \\
&= \frac{a - (N + 1)c_0 + Lc_1 + (N - L)c_0}{(N + 1)} \\
&= \frac{a - (L - 1)c_0 + Lc_1}{(N + 1)} \\
&= \frac{a - c_0 - L\Delta c}{(N + 1)}
\end{aligned}$$

The fifth line uses the definition $\Delta c = c_0 - c_1$. QED.

Lemma 3: *Interior solutions $L^* \in [0, N]$ satisfy the inequalities*

$$L^* \leq \frac{N + 2}{2} + \frac{a - c_0}{\Delta c} - \frac{(1 - \delta)\underline{k}(N + 1)^2}{2(\Delta c)^2} < L^* + 1$$

An implication of Lemma 3 is that the number of adopters can increase or decrease in N . This ambiguity arises because adding firms damps adoption incentives but increases the pool of possible adopters. The net effect depends on the parameter values.

Proof: Using Lemma 1, the inequalities (B.9) and (B.10) can be expressed

$$(1 - \delta)\underline{k} \leq [q(c_0, L^*; N) + \Delta c]^2 - q(c_0, L^* - 1; N)^2 \quad (\text{B.11})$$

$$(1 - \delta)\underline{k} > [q(c_0, L^* + 1; N) + \Delta c]^2 - q(c_0, L^*; N)^2 \quad (\text{B.12})$$

Next, applying the factoring relationship $x^2 - y^2 = (x + y)(x - y)$ provides

$$(1 - \delta)\underline{k} \leq (q_0(L^*) + \Delta c + q_0(L - 1))(q_0(L^*) + \Delta c - q_0^*(L^* - 1)) \quad (\text{B.13})$$

$$(1 - \delta)\underline{k} > (q_0(L^* + 1) + \Delta c + q_0(L^*))(q_0(L^* + 1) + \Delta c - q_0^*(L^*)) \quad (\text{B.14})$$

where we have suppressed the c_0 and N arguments for notational brevity. Focus on the top inequality. Substitute for $q_0(L)$ using Lemma 2. With some factoring, this yields

$$(1 - \delta)\underline{k} \leq \left(2\frac{a - c_0}{N + 1} - \left(\frac{L^*}{N + 1} + \frac{L^* - 1}{N + 1}\right)\Delta c + \Delta c\right) \left(-\frac{L^*}{N + 1}\Delta c + \frac{L^* - 1}{N + 1}\Delta c + \Delta c\right)$$

Next, collecting terms yields

$$(1 - \delta)\underline{k} \leq \left(2\frac{a - c_0}{N + 1} + \frac{N - 2L^* + 2}{N + 1}\Delta c\right) \left(\frac{N}{N + 1}\Delta c\right)$$

Move now to the bottom inequality (i.e., inequality (B.14)). The same manipulations yield

$$(1 - \delta)\underline{k} > \left(2\frac{a - c_0}{N + 1} + \frac{N - 2(L^* + 1) + 2}{N + 1}\Delta c\right) \left(\frac{N}{N + 1}\Delta c\right)$$

Combine the previous two inequalities, divide by $\frac{N}{N+1}\Delta c$, and subtract $2\frac{a-c_0}{\Delta c}$. This yields:

$$\frac{N - 2(L^* + 1) + 2}{N + 1}\Delta c < \frac{(1 - \delta)\underline{k}}{\Delta c} \frac{N + 1}{N} - 2\frac{a - c_0}{N + 1} \leq \frac{N - 2L^* + 2}{N + 1}\Delta c$$

Multiply by $N + 1$, divide by Δc , and subtract $N + 2$. This yields

$$-2(L^* + 1) < \frac{(1 - \delta)\underline{k}}{(\Delta c)^2} \frac{(N + 1)^2}{N} - 2\frac{a - c_0}{\Delta c} - (N + 2) \leq -2L^*$$

Finally, divide by negative two. This flips the direction of the inequalities and obtains the expression in Lemma 3:

$$L^* \leq \frac{N + 2}{2} + \frac{a - c_0}{\Delta c} - \frac{1}{2} \frac{(1 - \delta)\underline{k}}{(\Delta c)^2} \frac{(N + 1)^2}{N} < L^* + 1 \quad (\text{B.15})$$

QED.

Corollary: *Interior solutions $L^* \in [0, N]$ also satisfy the inequalities*

$$\frac{L^*}{N} \leq \frac{N + 2}{2N} + \frac{1}{N} \frac{a - c_0}{\Delta c} - \frac{(1 - \delta)\underline{k}}{2(\Delta c)^2} \frac{(N + 1)^2}{N^2} < \frac{L^* + 1}{N}$$

The corollary derives bounds for the fraction of firms that adopt in the absorbing state. It is sometimes more convenient to work with the fraction in the proof of proposition 1.

Proof of Proposition 1: The first part of the proposition states that $\underline{k} \leq 0$ implies $L^* = N$. This follows immediately from the assumptions of the model. Adoption happens (eventually) for any $L < N$ because $\pi(c_1, L; N) - \pi(c_0, L; N) \geq (1 - \delta)\underline{k}$ always if $\underline{k} \leq 0$, and this leads to $L^* = N$.

We turn now to the remainder of the proof. Denote the bound on the fraction of firms

that adopt in the absorbing states as $X(N)$. From the corollary to Lemma 3,

$$X(N) = \frac{N+2}{2N} + \frac{1}{N} \frac{a-c_0}{\Delta c} - \frac{(1-\delta)\underline{k}}{2(\Delta c)^2} \frac{(N+1)^2}{N^2}$$

By inspection, the limit of $X(N)$ is given by

$$\lim_{N \rightarrow \infty} X(N) = \frac{1}{2} - \frac{(1-\delta)\underline{k}}{2(\Delta c)^2}$$

This limit is less than half for any positive \underline{k} . Thus, there exists some n_1 such that if $N > n_1$ then $L^*/N < \frac{1}{2} < 1$ and $L^* < N$. Further, if $(1-\delta)\underline{k} > (\Delta c)^2$ then the limit is negative, and there exists some $n_2 > n_1$ such that if $N > n_2$ then $L^*/N = L^* = 0$. It remains to be shown that if $N > n_1$ then L^* weakly decreases in N . Here it is simpler to work with the bound on L^* , which we denote $Y(N)$. From Lemma 3,

$$Y(N) = \frac{N+2}{2N} + \frac{a-c_0}{\Delta c} - \frac{(1-\delta)\underline{k}}{2(\Delta c)^2} \frac{(N+1)^2}{N}$$

Taking a derivative yields

$$\frac{\partial Y(N)}{\partial N} = -\frac{1}{N^2} - \frac{1}{2} \frac{(1-\delta)\underline{k}}{2(\Delta c)^2} \frac{(N+1)^2}{N}$$

which is negative for any positive \underline{k} . Thus, L^* at least weakly decreases in N for any $N > n_1$ if $\underline{k} > 0$. QED.

C Additional Figures and Tables

Table C.1: Parameterizations used for Numerical Simulations

Parameter	Figure 1		Figure 2	Figure 3		Figure 4	Figure 5
	Top	Bottom		Top	Bottom		
a	2.50	3.72	4.00	4.12	3.00	30	3.72
c_0	1.00	0.31	1.00	1.60	1.97	1.5	0.31
c_1	0.70	0.21	0.90-0.50	1.06	1.77	0.5	0.21
\bar{k}	5.00	5.00	3.00	5.00	5.00	.	5.00
\underline{k}	1.00	1.93	0.00	3.10	1.21	40	1.93
δ	0.95	0.95	0.95	0.95	0.95	0.90	0.95
N	5	5	1-5	3-5	3-5	Many	3-5

Notes: The table provides the parameterizations used in Section 2. Adoption costs are assumed to have the uniform distribution.

Table C.2: Probit Regression Results with Interaction

Regressor	Equation for Shut Down vs. Maintain			
	(i)	(ii)	(iii)	(iv)
<i>Fuel Costs, Competition, and Demand</i>				
Fuel Costs	0.009 (0.010)	0.006 (0.010)	-0.022* (0.013)	-0.032*** (0.010)
Competitors	0.008 (0.013)	0.004 (0.015)	-0.015 (0.015)	-0.026** (0.012)
Fuel Costs × Competitors	0.000 (0.000)	0.000 (0.000)	0.001*** (0.000)	0.001*** (0.000)
Construction	-0.027** (0.011)	-0.026** (0.012)	-0.018 (0.012)	-0.005 (0.009)
<i>Mean Elasticities of Pr(Shut Down)</i>				
WRT Fuel Costs	0.52**	0.50**	0.10	-0.25
WRT Competitors	0.59**	0.54*	0.45	0.24
WRT Construction	-0.73***	-0.70**	-0.51*	-0.16
<i>Specification Details</i>				
Time Polynomial	no	1st Order	5th Order	no
Year Fixed Effects	no	no	no	yes

Notes: The table summarizes results obtained from multinomial probit regressions in columns (i)-(iii) and a binomial probit regressions in column (iv). The sample is comprised of 8,192 kiln-year observations over 1973-2013. The excluded instrument in the first stage is a 20-year lag on the number of nearby competitors. The elasticities are calculated for each observation and summarized with the mean. Standard errors are shown in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted with *, **, and ***, respectively.