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

Jeffrey Macher†
Georgetown University

Nathan H. Miller‡
Georgetown University

Matthew Osborne§
University of Toronto

October 14, 2016

Abstract

We estimate the effect of competition on the adoption of a cost-reducing technology in the cement industry, using data that span 1949-2013. The new technology, the precalciner kiln, reduces fuel usage and hence fuel costs. We find adoption is more likely if the cost savings are large, and less likely if there are many nearby competitors (accounting for the endogeneity of competition). We also find that competition damps the positive effect of cost savings. The results are consistent with theoretical models in which competition deprives firms of the scale necessary to recoup sunk adoption costs. We develop implications for environmental and antitrust policy.

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, Rich Sweeney, Mihkel Tombak, Francis Vella, and seminar participants at Georgetown University, Harvard Business School, and University of Toronto for helpful comments. We have benefited from conversations with Hendrick van Oss of the USGS and a number of industry participants.

†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 St., Toronto, ON, Canada M5S 3E6. Email: matthew.osborne@rotman.utoronto.ca
1 Introduction

The effect of competition on innovation has been a focus in the economics literature since at least the seminal work of Joseph Schumpeter (1934, 1942), which posits that large firms in concentrated markets invest more intensely in innovation. Empirical research on this subject is difficult because innovation is hard to measure and market concentration itself is endogenous. These concerns, however, have not deterred researchers: a vast number of articles examine descriptive regressions based on cross-industry comparisons of R&D spending and patent counts. Other contributions examine how competition and firm size affect technology adoption decisions (e.g., Oster (1982); Hannan and McDowell (1984); Rose and Joskow (1990); Schmidt-Dengler (2006)). More recently, a number of articles employ dynamic structural models to simulate the effect of competition, using theory to supplement empirical variation (e.g., Goettler and Gordon (2011); Igami (2015); Fowlie, Reguant and Ryan (2016); Igami and Uetake (2016)).

We study the diffusion of precalciner kiln technology in the United States portland cement industry. Several factors make this setting amenable to empirical analysis. In data spanning 1949-2013, we have annual observations on hundreds of older kilns that are candidates to be replaced with precalciner technology. The first adoption occurs in 1974, and precalciner kilns account for the bulk of domestic capacity by the end of the sample. Most older kilns that are not replaced with precalciner technology are shut down instead. The benefits of adoption derive from reduced production costs, and can be quantified with information on fossil fuel prices and the energy efficiency of old and new technology. Changes in fossil fuel prices over the sample period create exogenous variation in the cost savings available. The high transportation costs for portland cement make competitive conditions localized and plant-specific, so the benefits of adoption experienced by plants are heterogeneous. The institutional details of the industry suggest an instrument that provides plausibly exogenous variation in local competition.

To frame the analysis, we first extend the theoretical model of Dasgupta and Stiglitz (1980) on cost-reducing investments in Cournot equilibrium. This model incorporates precalciner adoption as both non-drastic and non-divisible: non-drastic because the fuel cost savings are insufficient to price competitors out of the market; and non-divisible because the decision is fundamentally binary. Within this context, we show that competition lessens

---

1 Aghion and Tirole (1994) refer to the Schumpeterian hypothesis regarding the impact of firm size and market structure on innovation as the second most tested relationship in industrial organization. Even the literature reviews are daunting (e.g., Kanien and Schwartz (1982); Baldwin and Scott (1987); Cohen and Levin (1989); Cohen (1995); Gilbert (2006); Cohen (2010)).
the adoption incentives of a myopic “focal firm” under mild conditions. We also solve a two-stage version of the model in which every firm can adopt the new technology, and show that competition limits adoption in subgame perfect equilibrium if investment costs are sufficiently large. The mechanism behind both results is deceptively simple: competition denies firms the scale required to recoup the investment costs of technology adoption. The model provides a number of other testable predictions that we take to the data.

The empirical model is based on a static game of imperfect information in which producers with non-precalciner kilns make a technology decision: (i) adopt precalciner technology; (ii) operate the older kiln; or (iii) shut down the older kiln without replacement. We parameterize the net benefits of adoption as depending on the number of nearby competitors, the fuel cost savings achieved with adoption, proximate construction activity, and various other controls. Theory suggests that the competition measure is positively correlated with the structural error term, which captures favorable but unobserved demand and cost conditions. This outcome introduces an endogeneity concern: to the extent that competition deters adoption, estimation risks understating the magnitude of this effect (or even suggesting that it is positive). We address endogeneity by using a 20-year competition lag as an excluded instrument, which is possible due to the length of the panel data. This instrument has power because kilns are long-lived, and is valid provided that autocorrelation in the structural errors is not too great. We estimate the model as a multinomial probit, using the two-stage conditional maximum likelihood estimator of Rivers and Vuong (1988).

The econometric results are consistent with the motivating theory. The presence of many nearby competitors diminishes the net benefits of adoption, and this effect is statistically significant, robust, and large: the mean elasticity of the adoption probability with respect to competition ranges between $-1.45$ and $-2.16$ in the baseline specifications. By contrast, the impacts of fuel cost savings and construction activity on technology adoption are positive, robust, and large: the mean elasticity of adoption with respect to fuel cost savings ranges from 0.57 and 0.82 and with respect to construction activity ranges from 1.16 and 1.54 in the baseline specifications. These magnitudes provide support for mechanism of the motivating theory: if demand and competition both increase by a similar amount then the net effect on adoption incentives is small. We also consider technology abandonment (i.e., kiln shutdowns without replacement). Our empirical results indicate that the probability of kiln shutdown increases with fuel costs and the number of nearby competitors, and decreases with proximate construction activity.

Our research has direct policy relevance in at least two arenas. First, our research is relevant to whether market-based CO₂ regulation would induce firms to adopt more efficient
“green” technology. Existing empirical articles on induced innovation typically find some margin of adjustment but do not address issues of competition (e.g., Newell, Jaffe and Stavins (1999); Popp (2002); Linn (2008); Aghion et al (2012); Hanlon (2014)). Our results indicate that firms are responsive if competition is not too great. This interactive effect arises both in the motivating theory and in the data. Our estimates imply that a monopolist facing a one standard deviation fuel cost shock is nearly five times more likely to adopt precalciner technology in response, relative to a firm facing an average number of nearby competitors. Firms with many nearby competitors are more likely to shut down than to adopt new technology. These findings raise interesting and important questions about dynamic adjustment paths that require a more sophisticated methodology than in this paper.

Second, our research is relevant to the antitrust review of mergers. Allegations that mergers among competitors 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, because the court documents typically do not elaborate on the theoretical mechanism by which market structure affects innovation incentives. Our results suggest one specific setting in which mergers could have pro-competitive effects on innovation; for innovations that are non-drastic and non-divisible, it is possible that consolidation allows firms to achieve the scale required to profitably recoup the fixed costs of investment. Our empirical results do not inform the appropriate standard under which such an efficiency should be deemed sufficiently substantial or cognizable.

The external validity of our results is best developed via a brief literature review. It has long been understood that market power can facilitate innovation (e.g., Dasgupta and Stiglitz (1980)). Yet the opposite effect can arise if, for example, innovation cannibalizes monopoly profit (Arrow (1962)), preemptive investments deter entry (Gilbert and Newbery (1982)), or firms innovate to escape competitive pressure (Aghion et al (2005)). As many other possibilities exist, we refer interested readers to Aghion and Griffith (2005) and Gilbert (2006) for useful and complementary literature reviews. The institutional details of the market matter because there is no single theory that applies to all situations. We believe the two most important institutional details in our application are that technology adoption is both non-drastic and non-divisible. These conditions enable the existence of competitors to deprive firms of the scale necessary to recoup investment costs.

With non-drastic and non-divisible investments, it is still possible for competition to speed adoption due to preemption incentives in the medium run, even as it limits adoption
in the long run. We do not find support for preemption in the data – the presence of early adopters has little additional explanatory power over adoption decisions. We suspect that this result may be due to the large number of competitors that the average plant faces, which thereby distinguishes cement from others settings in which researchers find evidence of preemption that typically feature either tight oligopolies (e.g., Genesove (1999); Vogt (2000); Schmidt-Dengler (2006); Gil, Houde and Takahashi (2015)) or a monopolist that invests to deter entry (e.g., Dafny (2005); Ellison and Ellison (2011)).

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 adoption has been found in a number of settings, including ATMs and credit scoring in banking (Hannan and McDowell (1984); Akhavein, Frame and White (2005)), 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)). This is consistent with the mechanism of the motivating theory, in which competition disciplines firm size. There are, of course, counter-examples: perhaps most notably, the case of the basic oxygen furnace in the steel industry (e.g., Oster (1982)). Within this literature, our research is distinguished by the amount of cross-sectional and time-series variation in the cement data, as well as by the availability of an instrument that provides plausibly exogenous variation in the competitive environment.

Among the recent dynamic structural articles on competition and innovation, the most directly relevant for our research is Fowlie, Reguant and Ryan (2016). These authors simulate the effects of market-based CO₂ regulation on the portland cement industry, based on a model that allows plants to make forward-looking capacity and exit decisions. State-space payoffs are determined by Nash-Cournot competition within local markets. The simulations indicate that regulation induces exit and capacity-reductions, which comports with our econometric results that higher fuel prices increase the propensity of plants to shut down older kilns. The authors’ simulation does not allow for technology adoption, however, which our results suggest is a meaningful margin of adjustment. We therefore view the two research projects as complementary. Overall, our empirical results support the usefulness of the underlying structural framework, and our theoretical extension of Dasgupta and Stiglitz (1980) helps clarify the mechanisms through which regulation affects dynamic decisions.

The paper proceeds as follows. Section 2 develops the motivating theory, including
both the focal firm model and the two-stage model. Section 3 provides institutional details on precalciner kilns and the portland cement industry. Section 4 develops the two-stage game of perfect information that we take to the data, and also discusses identification. 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 Models

We develop three complementary models that illuminate the relationship between competition and technology adoption. Each focuses on non-exclusive and non-drastic technologies that reduce marginal costs of production. We summarize the models here:

1. The “focal firm” model considers the adoption decision of a single myopic firm that subsequently competes a la Nash-Cournot. Competition reduces adoption incentives of the focal firm under weak conditions, and the model helps clarify the mechanisms at play. It also informs the empirical specification.

2. The “long run” model is a two stage game. Each firm first chooses whether adopt a cost-reducing technology, and then competes a la Nash-Cournot. The frequency of adoption decreases as the number of firms grows large, eventually reaching zero. This supports that competition can have long-lasting effects on technology adoption.

3. The final model is a simplified version of the dynamic oligopoly games used in theoretical and computational research to understand firm investment, entry, and exit decisions under uncertainty (e.g., Ericson and Pakes (1995); Doraszelski and Satterthwaite (2010)). The model incorporates incentives for preemption and thus the possibility that competition could spur adoption. If adoption is unlikely, however, preemption incentives are weak and the dynamic model tracks the focal firm model. Further, a grid search over the structural parameters supports that competition typically deters adoption if payoffs are determined by Nash-Cournot competition.

We focus on Nash-Cournot competition because it is consistent with non-drastic technology adoption: a modest reduction in one firm’s marginal costs does not eliminate the output of competitors (at least in the short run). Preemption incentives are stronger for drastic technology adoption. We develop this in the dynamic model by contrasting results under Nash-Cournot and Nash-Bertrand payoffs.
industry (e.g., Ryan (2012); Fowlie, Reguant and Ryan (2016)). The closest antecedent to our theoretical research is Dasgupta and Stiglitz (1980), which examines incremental cost-reducing investments in symmetric Nash-Cournot equilibrium and finds that R&D intensity is a decreasing function of the number of firms. Our models extend this analysis to the case of discrete cost-reducing technology adoption and clarify the mechanisms involved.

2.1 Focal firm model

Consider competition in a market with \( j = 1, \ldots, N \) firms that have constant marginal cost functions. A single “focal firm” can purchase a technology that lowers its marginal cost by paying the capital cost \( k \). The focal firm’s marginal costs with and without technology adoption are \( c_i^1 \) and \( c_i^0 \), respectively, such that \( c_i^1 = c_i^0 - \Delta c \geq 0 \). The equilibrium producer surplus of the focal firm is given by the function \( \pi_i(c_i, c_{-i}, \gamma) \), where \( c_{-i} \) is a vector of competitors’ marginal costs and \( \gamma \) incorporates all other factors. The focal firm adopts the technology if the benefit exceeds the cost:

\[
b_i(c_i^0, c_i^1, c_{-i}, \gamma) \equiv \pi_i(c_i^1, c_{-i}, \gamma) - \pi_i(c_i^0, c_{-i}, \gamma) > k
\]

**Assumption A1:** The equilibrium producer surplus function is differentiable.

A1 rules out undifferentiated Nash-Bertrand competition, but holds in most other standard oligopoly models. Yet even with only this weak assumption, it is still possible to derive a result of empirical relevance. Consider a hypothetical per-unit production subsidy, \( s \), that applies uniquely to the focal firm. The benefits of technology adoption can be reexpressed using a first order Taylor Series expansion:

\[
b(c_i^0, c_i^1, c_{-i}, \gamma) \approx \frac{\partial \pi_i(c_i - s, c_{-i}, \gamma)}{\partial s} \bigg|_{c_i = \hat{c}_i} \Delta c
\]

where \( \hat{c}_i = \frac{2c_i^0 - \Delta c}{2} \). To a first order approximation, competition and the other factors summarized in \( \gamma \) affect the benefits of technology adoption through their interaction with \( \Delta c \).

**Assumption A2:** Firms produce a homogeneous product and competition is Nash-Cournot. Prices are given by the inverse demand curve \( P(Q) = a - Q \), for \( Q = \sum_{j=1}^{N} q_j \).

The linearity assumption on demand yields tractable expressions for equilibrium markups
and quantities:

\[ P^*(c_j, \bar{c}, a, N) - c_j = q^*(c_j, \bar{c}, a, N) = \frac{a - c_j + N(\bar{c} - c_j)}{(N+1)} \]  

(3)

where \( \bar{c} = \frac{1}{N} \sum_{j \neq i} c_j \). The first equality implies that the equilibrium surplus function is given by \( \pi_j^*(c_j, \bar{c}, a, N) = (q^*(c_j, \bar{c}, a, N))^2 \). We restrict attention to markets in which quantities are positive for all firms. The unit slope normalization is without loss of generality, as all results derived below extend easily to demand curve rotations. We 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. The benefit that focal firm \( i \) receives from technology adoption is given by:

\[ b_i(\Delta_c, c_i^0, \bar{c}, a, N) = \frac{2N}{(N+1)} q^*(\hat{c}_i, \bar{c}, a, N) \Delta_c \]  

(4)

where \( \bar{c} = \frac{1}{N}(\sum_{j \neq i} c_j^0 + \hat{c}_i) \) is the average marginal cost evaluated at the midpoint between the focal firm’s new and old cost. The equation can be derived directly from the equilibrium markups and quantities.\(^4\) The following comparative statics are straight-forward:

**Result 1:** Under A1 and A2, the benefits of technology adoption:

(a) increase with the cost savings: \( \frac{\partial b_i(\Delta c, c_i^0, \bar{c}, a, N)}{\partial \Delta c} > 0 \)

(b) decrease with the focal firm’s cost: \( \frac{\partial b_i(\Delta c, c_i^0, \bar{c}, a, N)}{\partial c_i^0} < 0 \)

(c) increase with average costs: \( \frac{\partial b_i(\Delta c, c_i^0, \bar{c}, a, N)}{\partial \bar{c}} > 0 \)

(d) increase with demand: \( \frac{\partial b_i(\Delta c, c_i^0, \bar{c}, a, N)}{\partial a} > 0 \)

To assess how the level of competition (i.e., \( N \)) affects the benefits of adoption, it is necessary to specify the marginal costs of the firms being added or removed from the market. Let the marginal costs of these firms be \( \bar{c} \), so that average industry costs is unaffected. Then adding an addition firm decreases the benefit that the focal firm receives from adoption if the following condition holds:

\[ \left( \frac{\bar{c} - \hat{c}_i}{\hat{c}_i} \right) < \left( \frac{a - \hat{c}_i}{\hat{c}_i} \right) \left[ \frac{(N+1)^3 - N(N+2)^2}{N^2(N+2)^2 - (N+1)^4} \right] \]  

(5)

\(^4\)Alternatively, because the producer surplus of the focal firm is quadratic in its costs, the Taylor series approximation in equation (2) holds with equality, and equation (4) can be obtained by differentiating the producer surplus function with respect to the magnitude of a per-unit production subsidy.
The condition rules out that the focal firm has much lower costs relative to its competitors (see Appendix B for the derivation). It is likely to hold in the empirical application. Based on details that we describe later, reasonable estimates place the LHS of in the range of 0.15-0.20. The results of Ganapati, Shapiro and Walker (2016) can be manipulated to obtain \((p - c)/c = 1.50\) for the cement industry, and this provides a lower bound to \((a - \bar{c}_i)/\bar{c}_i\). Lastly, the term in brackets equals 0.294 if \(N = 2\), and converges quickly to 1/2 as \(N\) grows large. This leads to the second set of theoretical results:

Result 2: Under A1, A2, and condition (5), the benefits of technology adoption decrease in the number of competitors: 
\[
\frac{\Delta b_i(\Delta c, c_0, a, \bar{c}, N)}{\Delta N} < 0.
\]

Also, increasing \(N\) reduces the (positive) effect of \(\Delta c\) on adoption benefits: 
\[
\frac{\Delta}{\Delta N} \frac{\partial b_i(\Delta c, c_0, a, \bar{c}, N)}{\partial \Delta c} < 0.
\]

Finally, to build intuition on why the benefit of technology adoption decreases with \(N\), reconsider the approximation in equation (2). Under Nash-Cournot competition but relaxing the linearity demand assumption, the derivative of the producer surplus function can be decomposed as follows:

\[
\frac{\partial \pi_i(c_i - s_i, c_{-i}, \gamma)}{\partial s} = q^*_i(c_i, c_{-i}, \gamma) - \sum_{k \neq i} \frac{\partial P^*(Q)}{\partial q_k} \frac{\partial q^*_k(c_k, c_{-k}, \gamma)}{\partial q_i} \frac{\partial q^*_i(c_i, c_{-i}, \gamma)}{\partial s}
\]

The first term (a cost savings effect) represents how lower marginal cost increases the variable profit for each unit produced (Arrow (1962)). This effect provides the main mechanism through which increasing \(N\) reduces the adoption incentive: more competitors means lower equilibrium output, which reduces the available cost savings. The second term (a strategic effect) represents how the focal firm – given its lower marginal costs – induces competitors to produce less, which subsequently raises price. The strategic effect simplifies to 
\[
q^*_i(c_i, c_{-i}, a, N) \frac{N-1}{N+1}
\]

with linear demand; thus, increasing \(N\) reduces the cost saving effect but can amplify the strategic effect: condition (5) determines which change dominates.

2.2 Long run model

We now extend the motivating theory to a setting in which all firms have the opportunity to adopt a cost reducing technology. Specifically, we examine a game with two stages. In the

\[\text{Precalciner technology reduces fuel costs by about 30 percent. Suppose there are no other costs (a conservative assumption). Then the LHS is about 0.176. The findings of Ganapati, Shapiro and Walker (2016) are based on data from the Census Bureau’s Census of Manufactures. Table 2 of the May 2016 draft indicates prices of 0.05 and costs of 0.02 (in thousands of 1987 dollars per cubic yard).}\]
first stage, each of \(N\) firms can adopt a technology by paying the capital cost \(k\). In the second stage, all firms compete according to undifferentiated-products Nash-Cournot. Marginal costs with and without the technology are \(c_1\) and \(c_0\), respectively, and again \(c_1 = c_0 - \Delta c \geq 0\). Prices are given by the inverse demand curve \(P(Q) = a - Q\), for \(Q = \sum_{j=1}^{N} q_j\).

We characterize the number of firms that adopt the technology in subgame perfect equilibrium (SPE). We first summarize second stage payoffs. The stage-game equilibrium markups and quantities of a non-adopter are:

\[
P^*(L) - c^0 = q^*(c^0, L) = \frac{a - L\Delta c - c_0}{N + 1}
\]

where we use \(L\) to denote the number of adopters. This result is obtained by manipulating equation (3). The solutions for adopters can be shown to be given by \(q^*(c_1, L) = q^*(c_0, L) + \Delta c\). Adopters thus have greater quantity and larger markups, and both effects increase with the cost savings available with the new technology.

We next turn to the first stage. The SPE is characterized by some number of adopters, \(L^* \leq N\), such that (i) all adopters prefer adoption to non-adoption, and (ii) all non-adopters prefer non-adoption to adoption. Recalling that producer surplus in the second stage equals the square of equilibrium quantity, these stability conditions generate two inequalities that can be manipulated to solve the game:

\[
k \leq q^*(c_1, L = L^*)^2 - q^*(c_0, L = L^* - 1)^2
\]

and

\[
k > q^*(c_1|L = L^* + 1)^2 - q^*(c_0|L = L^*)^2
\]

Plugging in for the equilibrium quantities, the following expression obtains with some algebra that we defer to Appendix B.2:

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

By inspection, the SPE number of adopters can increase or decrease in \(N\). This ambiguity arises because more firms simultaneously damps adoption incentives but increases the pool of possible adopters: the net effect depends on the parameter values. If capital costs are high enough relative to the cost savings (specifically, if \(k > 2(\Delta c)^2\)), however, then the SPE number of adopters approaches zero as the number of firms grows large. Also relevant here is the fraction of firms that adopt in the SPE, which can be obtained by dividing equation
This fraction is less than one for sufficiently high $N$, and approaches zero as $N$ grows large under the same condition that $k > 2(\Delta c)^2$. Figure 1 plots the number and fraction of adopters under one such parameterization ($a = 20, k = 4, \Delta c = 1$). For $N < 7$, the number of adopters grows with the number of firms because all firms find it profitable to adopt. The number of adopters then shrinks for $N > 7$, and equals zero for $N \geq 14$.

**Result 3:** Under the condition $k > 2(\Delta c)^2$, there exists some $N_1$ and $N_2$ ($N_1 < N_2$) such that (1) if $N > N_1$ then the fraction of firms that adopt the technology in SPE decreases with $N$ and (2) if $N > N_2$ then no firms adopt.

## 2.3 Dynamic model

### 2.3.1 Framework, policies, and equilibrium

Again consider a market with $j = 1, \ldots, N$ firms that can adopt a cost-reducing technology. Marginal costs equal $c_1$ and $c_0$ with and without the technology, respectively. In each period, each firm $i$ that has not (yet) adopted the technology receives a private draw on adoption costs, $k_i$, that is drawn from a continuous distribution $F(\cdot)$ with support $[k, \overline{k}]$. It decides whether to adopt and then plays the stage game, which is characterized by static competition.
in quantities or price. Adoption is irreversible in subsequent periods. We characterize firm behavior in a symmetric Markov-perfect equilibrium in pure strategies.

There is a single state variable that governs adoption decisions in period $t$: the number of competitors that adopted prior to period $t$. We denote the state variable as $L_t$. Because there are $N - 1$ competitors, the state variable can take on values from 0 to $N - 1$. The firms’ actions are to adopt, which we index as 1, or to not adopt, which we index as 0. Static profit from each action $a \in \{0, 1\}$ is $\pi_a(L_t; N)$. The following assumption is standard and helps ensure that a symmetric equilibrium in pure strategies exists:

**Assumption A3:** (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., $\pi_a(L_t; N)$ is the same for all firms $i$.

Denote the value functions heading into period $t$, i.e., before the cost draws are received, as $V_0(L_t; N)$ and $V_1(L_t; N)$. 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 $v_a(\cdot)$ denotes the expected discounted profit for action $a \in \{0, 1\}$. 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 $L_{t+1} = 0, 1, \ldots, N - 1$ 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.

It is also helpful to write the flow profits and value functions in vector form. For actions $a \in \{0, 1\}$, let $\mathbf{\pi}_a(N) = (\pi_a(0; N), \pi_a(1; N), \ldots, \pi_a(N - 1; N))'$ and $\mathbf{V}_a(N) = (V_a(0; N), V_a(1; N), \ldots, V_a(N - 1; N))'$.

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

$$
\text{Upgrade: } v_1(L_t; N, k) = \mathbf{P}_0(L_t; N)'(\mathbf{\pi}_1(N) + \delta\mathbf{V}_1(N)) - k, \tag{11}
$$

$$
\text{Not Upgrade: } v_0(L_t; N) = \mathbf{P}_0(L_t; N)'(\mathbf{\pi}_0(N) + \delta\mathbf{V}_0(N))
$$

---

6The support of the adoption cost distribution can be bounded or unbounded, i.e., it can be the case that $k = -\infty$ or $k = \infty$. A bounded distribution is theoretically attractive because if $k \geq 0$ 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.

7If $L_t = 0$ 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 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 $P(L_t; N) = F(k^*(L_t; N))$.

The value function associated with adoption, $V_1(L_t; N)$, has an explicit solution because adoption is irreversible. Define the upper triangular matrix $\Pi_0$ as follows:

$$\Pi_0 = \begin{bmatrix} P_0(0; N)' \\ P_0(1; N)' \\ \vdots \\ P_0(N - 1; N) \end{bmatrix}. \quad (12)$$

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 $\Pi_1$ characterize the post-adoption transitions of competitors. Then the value functions associated with adoption are given by the vector

$$V_1(N) = \Pi_0 (I + \delta(I - \delta\Pi_1)^{-1}) \pi_1(N). \quad (13)$$

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

$$V_0(L_t; N) = \int_{k}^{k_{\text{max}}} \max\{v_0(L_t; N), v_1(L_t; N, k)\} dF(k). \quad (14)$$

**Assumption A4:** Define an industry state transition matrix, $\tilde{\Pi}$, 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 A3 and A4, Propositions 2 and 5 of Doraszelski and Satterthwaite (2010) guarantees the existence of a symmetric pure strategy Nash equilibrium. This helps motivates our empirical model, which assumes that firms are symmetric and take the same actions for any values of the observables and unobservables. The dynamic game is simpler than that of Doraszelski and Satterthwaite (2010) because investment is not continuous and exit

---

8The matrix $\Pi_1$ is composed of stacked vectors of post-adoption transition probabilities $P_1(L_t; N)$. Because an adoption changes all subsequent adoption probabilities, $P_1(L_t; N)$ is different than $P_0(L_t; N)$ For example, if $N = 2$ then $P_1(0; N) = (1 - P(1), P(1))$, but $P_1(1; N) = (0, 1)$.
is prohibited. These changes do not materially affect the proofs and it is possible to apply standard dynamic programming arguments, as well as Brouwer’s fixed point theorem, following exactly the arguments in Doraszelski and Satterthwaite (2010).\footnote{The 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 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.}

2.3.2 Benefits of adoption

As with the focal firm model, some insights can be gained by approximating the benefits of adoption. Define these as 

\[ b(L_t; N, k) = v_1(L_t; N, k) - v_0(L_t; N), \]

and let \( b = (b(0; N, k), b(1; N, k), \ldots, b(N - 1; N, k))' \). After a series of intermediate calculations provided in the appendix, an approximation to the benefit of upgrading is

\[
b \approx \Pi_0 \begin{pmatrix} \frac{\partial \pi}{\partial s} \Delta c + \\ \text{static part} \end{pmatrix} \delta \left( \left[ \Pi_0 \left( I + \delta (I - \delta \Pi_1)^{-1} \right) - A - B \right] \pi_0 + \left( \Pi_0 \left( I + \delta (I - \delta \Pi_1)^{-1} \right) - A \right) \frac{\partial \pi}{\partial s} \Delta c - C \right),
\]

where

\[
A = \left( I - \delta \Pi_0 (I - F) \right)^{-1} \Pi_0 F \left( I + \delta \Pi_0 \left( I + \delta (I - \delta \Pi_1)^{-1} \right) \right)
\]

\[
B = \left( I - \delta \Pi_0 (I - F) \right)^{-1} \left( I - F \right)
\]

\[
C = \left( I - \delta \Pi_0 (I - F) \right)^{-1} F \kappa
\]

\( F \) is a diagonal matrix with \( F(k^*(L)) \) on each \( L + 1 \)st diagonal element, \( \kappa \) is a diagonal matrix with \( E(k|k < k^*(L)) \) on each \((L + 1)^{\text{st}}\) diagonal element, and \( \frac{\partial \pi}{\partial s} \) is a vector of the derivatives of the profit function at each state with respect to the change in cost.\footnote{To obtain equation \( (15) \), one solves for both the value functions \( V_0 \) and \( V_1 \) in terms of the period profits \( \pi_0 \) and \( \pi_1 \) and the state transitions. The benefit of upgrading then can be expressed in terms of state transitions, upgrade probabilities, and per period profits. At this stage, one can substitute in the Taylor expansion of \( \pi_1 \) around \( c_0 \). The idea behind the approximation is therefore similar to that behind the approximation of the focal firm upgrade benefit in the static game, except for one difference: we do not directly take a Taylor approximation of the benefit, as that would involve differentiating the state transitions (which themselves are functions of per-period profits).}
We break equation (15) into two pieces: a static part and a dynamic part. The static part behaves in the same way as the static approximations in the focal firm model. The behavior of the dynamic part is harder to pin down because the state transitions $\Pi_0$ and $\Pi_1$ and the upgrade probabilities $F$ are complicated functions of the underlying model primitives. However, if adoption probabilities are small, which is the case in our empirical setting, then $\Pi_0$ and $\Pi_1$ are close to the identity matrix and $F$ is close to the zero matrix. As a result, the matrix $A$ will be close to zero, and the matrix $B$ will approach $\frac{1}{1-\delta}I$. This implies that the term which multiplies $\pi_0$ approaches zero. As a result, if adoption probabilities are small, it is the case that

$$b \approx \frac{1}{1-\delta} \frac{\partial \pi}{\partial s} \Delta c - \delta C,$$

and the implications derived for the static model hold for the dynamic model.

### 2.3.3 Numerical simulations

To be written.

## 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. Because the associated fuel costs account for a sizable portion of revenues, the fuel efficiency of the kiln technology employed is an important determinant of a plant’s overall profitability.

Plants equipped with precalciner technology preheat the raw materials using the exhaust gases of the kiln combined with heat from a supplementary combustion chamber. This approach reduces the energy requirements of production by 25-35 percent relative to older wet and long dry kilns. Because precalciners allows one of the main chemical reactions to occur before raw materials enter the kiln, the requisite kiln length is greatly reduced, and installation requires a retrofit of the entire plant. Cement producers typically outsource the design
Table 1: The Portland Cement Industry over 1973-2013

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

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.

to one of several industrial architectural firms with expertise handling space constraints and geological considerations. Installation then is completed by an industrial construction company. The physical component is not especially demanding—many industrial construction firms can manage the steel plates, refractory linings, and duct work—but the total cost of design and installation nonetheless is large. As a result, adoption of precalciner technology has been gradual since its development in the late 1960s and early 1970s.

Table 1 tracks precalciner kiln adoption over time. In 1973, nearly all plants used inefficient wet and long dry kilns. A small number used preheater technology, which recycles exhaust gases without a supplementary combustion chamber, but none used precalciners. Over the ensuing four decades, the number of wet kilns decreased from 249 to 19, the number of long dry kilns decreased from 157 to 26, and the number of precalciner kilns increased substantially. In the final year of data, 66 of the 140 kilns in operation use precalciners and account for 74 percent of industry capacity. Indeed, the higher capacity of precalciner kilns explains why industry capacity increased as the total number of plants and kilns decreased.

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
Table 2: Fuel Costs per Metric Tonne of Cement

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

Notes: The table provides average fuel costs by kiln technology, the hypothetical fuel costs of a kiln with “frontier technology” defined as a precalcer 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.

Firms would respond to the market-based regulation of CO₂ emissions (which would change the implicit price of fossil fuels).

Table 2 also provides the fuel costs of the “frontier technology,” which we define as a precalcer 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 precalcer 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. The 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 to 33 percent of revenues. Two recent papers estimate that pass-through of fuel costs to price in the 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 2 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 few.

11There 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 prices adjust accordingly (e.g., Busse, Knittel and Zettelmeyer (2013); Langer and Miller (2013)).
single nearby plant. These differences provide useful cross-sectional variation.

Cement is used in construction projects, so demand is highly procyclical. Figure 3 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 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 a snapshot of the industry, as of December 31, 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. The supplementary dataset is useful because it allows us to construct lagged competition 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. We use county-
level data on construction employment and building permits from the Census Bureau to account for demand-side fluctuations.\textsuperscript{14} 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\textsuperscript{15} The data for 1973-1985 are obtained from the University of Michigan Data Warehouse. The building permits data are maintained online by the U.S. Department of Housing and Urban Development.\textsuperscript{16} 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

We use multinomial probit regressions to examine the determinants of technology adoption and kiln shutdown. We interpret the regressions as approximating a static game of imperfect information in which producers with non-precalciner kilns make a technology decision and then compete in a stage game. Producers can be conceptualized as playing this static game each year, which exploits the panel data. Alternatively, the regressions could be interpreted as approximating the policy functions from a dynamic game, following the standard two-step method for estimating dynamic games developed in Bajari, Benkard and Levin (2007)\textsuperscript{17} The dual interpretations are unproblematic given the theoretical result of Section 2.3 that the dynamic game mimics the static game if transition probabilities are small.

4.1 Payoffs and policies

Producer surplus in the stage game depends on marginal costs, demand, the number of firms, control variables, and a set of private stochastic shocks. The deterministic portion of producer surplus is denoted $\pi(c_{it}, \bar{c}_{it}, N_{it}, a_{it}, w_{it}; \theta)$, where the first four arguments are defined as in the focal firm model (for $a = \{0, 1\}$), $w_{it}$ is a vector of controls, and $\theta$ is a vector of parameters. Producers make their technology choice to maximize profit. Profit in

\textsuperscript{14}For 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.


\textsuperscript{17}The first step of this method involves regressing the discrete firm choice on exogenous variables to recover the policy functions; the second step uses forward simulation to recover the dynamic structural parameters. We focus exclusively on the first step because our objective is to better understand firm policies, rather than to conduct counterfactual simulations.
the event of adoption is producer surplus less an investment cost, and profit in the event of shutdown is zero. We assume that the “benefit of adoption” can be written

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

which is consistent with the focal firm model. With this notation in hand, producers make their technology decision to maximize \( \Pi_{it} \), given by

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

(17)

where \((u_{it}^A, u_{it}^0, u_{it}^S)\) are the stochastic shocks. We parameterize the producer surplus functions using linear approximations. Thus, we do not seek to recovery any underlying structural parameters, instead focus on understanding the empirical determinants of technology choices. Define the scalars \( y_{it}^A \equiv b_{it}(\cdot) - k_{it} + u_{it}^A \), and \( y_{it}^S \equiv -\pi(c_{it}^0, \cdot) + u_{it}^S \). Under the linear approximations, we have

\[
y_{it}^A = \beta_1^A \Delta c_{it} + \beta_2^A N_{it} + \beta_3^A a_{it} + \beta_4^A \bar{c}_{it} + \beta_5^A c_{it}^0 + w_{it}' \alpha^A + \phi_t^A + u_{it}^A
\]

(18)

\[
y_{it}^S = \beta_1^S c_{it}^0 + \beta_2^S N_{it} + \beta_3^S a_{it} + \beta_4^S \bar{c}_{it} + w_{it}' \alpha^S + \phi_t^S + u_{it}^S
\]

(19)

We specify \( \phi_t^A \) and \( \phi_t^S \) alternately using linear time trends, a flexible polynomial in time, and year fixed effects. This helps accounts for any time-related changes, including learning-by-doing in precalciner installation. The control variables include kiln age and kiln capacity.

### 4.2 Estimation

Because the stochastic shocks summarize the net effect of unobserved demand and cost factors, there is a potentially confounding correlation between these shocks and the number of firms. 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 the following equation:

\[
N_{it} = z_{it} \gamma_1 + \Delta c_{it} \gamma_2 + a_{it} \gamma_3 + \bar{c}_{it} \gamma_4 + c_{it}^0 \gamma_5 + w_{it}' \gamma_6 + \phi_t^N + \nu_{it}
\]

(20)
where $z_{it}$ is an instrument that is excluded from the structural equations, $\phi_{it}^N$ is specified the same way as $\phi_{it}^A$ and $\phi_{it}^S$, and $v_{it}$ is a reduced-form error term. For notational convenience, 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:

$$\Omega \equiv \begin{bmatrix}
\sigma_{uu}^A & \sigma_{vu}^A & \sigma_{vv}^A \\
\sigma_{vu}^A & \sigma_{vv}^A & \sigma_{vu}^S \\
\sigma_{vv}^A & \sigma_{vu}^S & \sigma_{vv}^S 
\end{bmatrix}$$ (21)

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, however, equations (18) and (19) can be rewritten as:

$$y_{it}^A = \beta_1^A \Delta c_{it} + \beta_2^A N_{it} + \beta_3^A a_{it} + \beta_4^A \bar{c}_{it} + \theta^A_{it} + \phi_{it}^A + \eta_{it}^A$$ (22)

$$y_{it}^S = \beta_1^S c_{it}^0 + \beta_2^S N_{it} + \beta_3^S a_{it} + \beta_4^S \bar{c}_{it} + \theta^S_{it} + \phi_{it}^S + \eta_{it}^S$$ (23)

where $\lambda^k = \sigma_{vu}^k / \sigma_{vv}$ and $\eta_{it}^k = u_{it}^k - v_{it} \lambda^k$ for $k \in \{A, S\}$. 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. Use OLS to regress $N_{it}$ on the exogenous regressors. This obtains an estimate of the reduced-form error term that we denote $\hat{v}_{it}$.

2. Estimate the multinomial probit model of equations (22) and (23) with maximum likelihood, using $\hat{v}_{it}$ as a control function to account for endogeneity. This approach obtains consistent estimates of the structural parameters $(\beta^k, \alpha^k)$ for $k \in \{A, S\}$. Differences between $v_{it}$ and $\hat{v}_{it}$ are normally distributed and consistent with the distributional assumptions of the multinomial probit model. The estimates of $\lambda^A$ and $\lambda^S$ can be used to test for the exogeneity of the competition measure.

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 the steps described in 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 correction for autocorrelation, but this too has little effect on the magnitudes of the standard errors.
4.3 Identification and instrument

The focal firm model indicates that technology adoption is more likely under favorable profit conditions (e.g., high demand, low marginal costs, or low capital costs), and greater profit supports more competitors in standard models. 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.

Finding an instrument to correct endogeneity bias is not straightforward. Our setting differs from more standard applications in industrial organization that require the estimation of either demand or supply, and for which cost or demand shocks respectively are valid instruments. Both demand and cost variables enter in the profit function in our application so neither can be used as instruments. Our instrument is instead a lagged version of the competition measure, and if the structural error terms do not exhibit autocorrelation then this instrument is valid. This instrument has power in the first stage because kilns tend to operate for many decades. The potential problem is that unobserved profit shocks may be persistent, but we use such long lags (up to 20 years) that any inter-temporal correlations between the instrument and the structural error plausibly are quite weak. As a test, we use alternative instruments based on lags of 5, 10, and 15 years, and the results are consistent with the effects of autocorrelation dying out over the longer time horizons.

Other sources of endogeneity seem unlikely. Technology decisions within a market are not likely to drive demand, because cement represents a small fraction of total construction costs; hence, exogeneity of the demand controls is likely to be reasonable. 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, bituminous coal prices do not exhibit the same pro-cyclical variation as cement demand. Industry costs (i.e., $c_{it}$) 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 fired with the most affordable fuel. This difference provides the empirical proxy for the \(\Delta c\) term that appears in the motivating theory.

We measure competition based on plant locations and gasoline prices (which scale transportation costs). Define a distance metric as the multiplicative product of miles and a gasoline price index that equals one in the year 2000. We calculate the number of competing plants within a 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 more than 400 miles are unlikely to compete for customers.\(^{18}\) 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 of 400 to calculate the relevant notion of industry costs for each plant (i.e., \(c_{it}\)); this is defined as the average costs of plants with the radius. For both the competition and industry cost variables, use the location of plants/kilns as of the prior year so they are unaffected by the technology decisions in the current year.

Figure 4 provides histograms for the count of nearby competitors, separately for each decade in the data. Cross-sectional variation is due to the dispersion in plant locations, while inter-temporal variation arises due to gasoline price fluctuations and the net decrease in the

\(^{18}\)Our treatment of distance reflects the predominant role of trucking in 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)).
number of plants over the sample period. We use instruments based on the locations of plants 20 years prior to the observation 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. The instruments are constructed based on the plants in 1980 within 400 miles of the kiln’s location, even though the 1980 gasoline index differs than 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 first two controls are straight-forward and obtained from the PIS kiln data. The third 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. We also have constructed variables that capture the distance between plants and the nearest customs district through which foreign imports enter, but these controls have little explanatory power and we omit them from the specifications shown below.
Table 3: Number of Observations per Kiln

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Mean Obs.</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Kilns</td>
<td>460</td>
<td>17.81</td>
<td>2</td>
<td>6</td>
<td>12</td>
<td>34</td>
<td>41</td>
</tr>
<tr>
<td>Replaced Kilns</td>
<td>144</td>
<td>15.39</td>
<td>2</td>
<td>5</td>
<td>8</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>Shut Down Kilns</td>
<td>244</td>
<td>12.82</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td>Surviving Kilns</td>
<td>72</td>
<td>37.57</td>
<td>37</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
</tr>
</tbody>
</table>

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 provides information on the composition of the sample. 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. In total, there are 8,192 kiln-year observations in the regression 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: the mean of the indicators imply an empirical probability of 1.8% and 3.0%, respectively. The bivariate correlation coefficients show that there are limits to what can be identified given the available empirical variation. We use three normalizations to pair down the regressions specifications shown in equations (22) and (23):

1. We impose that $\beta_0^A = 0$ 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 ($\rho = 0.89$). This focuses the analysis on the effect of cost savings, which the motivating theory 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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Adoption</td>
<td>0.018</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Shutdown</td>
<td>0.030</td>
<td>0.17</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Fuel Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Cost Savings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Industry Costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Competitors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Construction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) Kiln Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) Kiln Capacity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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. 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. We therefore impose $\beta_4^A = 0$ in some regressions. 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 Main results

Table 5 presents the baseline probit results. Panel A addresses the likelihood of precalciner kiln adoption and Panel B addresses the likelihood of shutdown, both relative to the alternative of maintaining the older kiln. The columns control for changes over time in different ways: column (i) relies exclusively on the regressors; column (ii) adds a linear time trend.
\( t = 0, 1, \ldots, 40 \); 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 parameter estimates are consistent with the focal firm model of Cournot competition. Panel A shows that the benefits of adoption decrease with the number of nearby competitors but increase with the available cost savings, construction activity, and industry average costs. Each of the estimated coefficients are statistically significant in each specification (noting that the effect of industry costs is identifiable only in columns (i) and (ii)). To evaluate magnitudes, we calculate the elasticity of the adoption probability with respect to the variables of interest for each kiln-year observation in the data. The mean elasticity with respect to cost savings ranges over from 0.58 to 0.71; the mean elasticity with respect to the number of nearby competitors ranges from \(-1.41\) to \(-2.39\), and the mean elasticity with respect to nearby construction activity ranges from 1.13 to 1.89. We also highlight that the first stage residual that controls for unobserved (but favorable) demand and cost conditions is positive and statistically significant in Panel A, which again is consistent with the focal firm model and suggests that the IV strategy is important.

Panel B shows that relative profit of kiln shutdown increases in fuel costs and the number of nearby competitors, and decreases with construction activity. The precision of the coefficients is somewhat greater when controls for time are not too flexible. In particular, once time-series variation is removed, the remaining cross-sectional variation is insufficient to provide precise estimates of how fuel costs affect shutdown.

In the baseline regressions presented above, the effect of fossil fuel prices depends on competition due to the probit functional form. Consider that having many nearby competitors makes adoption sufficiently unlikely that there is little scope for other variables to matter. The focal firm model goes somewhat further: competition should damp how cost saving affect the benefit of technology adoption. We test this prediction by adding interactions to the specifications. Table 6 summarizes the results for the adoption decision. The interaction of cost savings and the nearby competitors is negative and statistically significant in columns (i)-(iii), which provides additional support for the theory. The level effect of competition retains its level effect in columns (i) and (ii), and this result dissipates with more flexible time controls. Neither it nor the interaction achieve statistical significance with year fixed effects. The mean elasticities are similar in magnitude to those of the baseline specifications. The coefficients obtained for the shutdown equation are shown in Appendix Table C.1. They take the expected signs but are not statistically significant independently (though some joint significance is obtained in columns (i) and (ii)).
<table>
<thead>
<tr>
<th>Regressor</th>
<th>Panel A: Adopt vs. Maintain</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fuel Costs, Competition, and Demand</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost Savings</td>
<td>$\Delta c_{it}$</td>
<td>0.051***</td>
<td>0.045***</td>
<td>0.040***</td>
<td>0.032***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Competitors</td>
<td>$N_{it}$</td>
<td>-0.055***</td>
<td>-0.058***</td>
<td>-0.037***</td>
<td>-0.026***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.009)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Construction</td>
<td>$a_{it}$</td>
<td>0.070***</td>
<td>0.078***</td>
<td>0.049***</td>
<td>0.036***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.012)</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Industry Costs</td>
<td>$\tau_{it}$</td>
<td>0.027***</td>
<td>0.029***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiln Age</td>
<td>$w_{1,it}$</td>
<td>0.018***</td>
<td>0.021***</td>
<td>0.019***</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Kiln Capacity</td>
<td>$w_{2,it}$</td>
<td>-0.805**</td>
<td>-0.675</td>
<td>-0.719</td>
<td>-0.418</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.392)</td>
<td>(0.445)</td>
<td>(0.447)</td>
<td>(0.341)</td>
</tr>
<tr>
<td>First Stage Residual</td>
<td>$\hat{v}_{it}$</td>
<td>0.075***</td>
<td>0.065***</td>
<td>0.083***</td>
<td>0.050***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.014)</td>
<td>(0.017)</td>
<td>(0.019)</td>
<td>(0.014)</td>
</tr>
<tr>
<td><strong>Mean Elasticities of Pr(Adoption)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRT Cost Savings</td>
<td></td>
<td>0.71</td>
<td>0.63</td>
<td>0.58</td>
<td>0.70</td>
</tr>
<tr>
<td>WRT Competitors</td>
<td></td>
<td>-2.25</td>
<td>-2.39</td>
<td>-1.57</td>
<td>-1.41</td>
</tr>
<tr>
<td>WRT Construction</td>
<td></td>
<td>1.79</td>
<td>1.89</td>
<td>1.28</td>
<td>1.13</td>
</tr>
<tr>
<td><strong>Specification Details</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time Polynomial</td>
<td>no</td>
<td>1st Order</td>
<td>5th Order</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>

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 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)

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

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 6: Probit Regression Results with Interaction

<table>
<thead>
<tr>
<th>Adopt Precalciner Technology vs. Maintain Old Kiln</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fuel Cost Savings, Competition, and Demand</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost Savings</td>
<td>0.066***</td>
<td>0.069***</td>
<td>0.066***</td>
<td>0.046***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.017)</td>
<td>(0.014)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Competitors</td>
<td>-0.040***</td>
<td>-0.039***</td>
<td>-0.014</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.013)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Cost Savings × Competitors</td>
<td>-0.0011*</td>
<td>-0.0015**</td>
<td>-0.0017**</td>
<td>-0.0009</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Construction</td>
<td>0.063***</td>
<td>0.066***</td>
<td>0.038***</td>
<td>0.031***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Industry Costs</td>
<td>0.027***</td>
<td>0.030***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Derived Statistics: Mean Elasticities of Pr(Adoption)**

<table>
<thead>
<tr>
<th></th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRT Cost Savings</td>
<td>0.59</td>
<td>0.54</td>
<td>0.47</td>
<td>0.55</td>
</tr>
<tr>
<td>WRT Competitors</td>
<td>-1.98</td>
<td>-2.03</td>
<td>-1.11</td>
<td>-1.18</td>
</tr>
<tr>
<td>WRT Construction</td>
<td>1.61</td>
<td>1.67</td>
<td>0.98</td>
<td>0.99</td>
</tr>
</tbody>
</table>

**Specification Details**

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Polynomial</td>
<td>no</td>
<td>1st Order</td>
<td>5th Order</td>
<td>no</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

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.
The augmented specification allows us to explore how competition mediates the effects of higher fossil fuel prices on adoption and shutdown. Figure 5 considers one standard deviation increases in cost savings and fuel costs. 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 three percentage points. This is large because the unconditional probability of adoption is 1.8 percent. By contrast, the effect for 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. Again the magnitudes are large relative to the unconditional probability of shutdown 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.

6.2 Preemption incentives and the inverted-U

The dynamic model of Section 2.3 indicates that preemption incentives are weak if the probability of precalciner adoption is small, as it is in our data. We test this prediction based
on the logic that if such preemption incentives are important then the presence of nearby precalciners should discourage adoption and/or encourage shutdown. To implement, we add the number of precalciner competitors within a radius of 400 to the baseline specifications. Two first stage regressions are required with the specification 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). Both first-stage residuals are used as controls in the second-stage regressions. Table 7 provides the results for adoption in column (i) and shutdown in column (iii), using a fifth-order polynomial to control for time effects. As shown, the number of nearby competitors retains its negative effect on precalciner adoption and its positive effect on shutdown. The coefficients on nearby precalciner competitors are smaller and quite imprecisely estimated. Similar results are obtained for other specifications of the time effects. Thus, the regression does not provide support for preemption being important in the cement industry.

The focal firm model does not predict an inverted-U relationship in which adoption incentives are maximized with intermediate levels of competition. Instead, competition deprives firms of the scale required to recoup investment costs, and this effect is monotonic. The inverted-U relationship arises in some other models (e.g., Aghion et al (2005)), however, and we test for the relationship by allowing the benefits of adoption to have a quadratic relationship with the competition measure. The results are provided in columns (ii) and (iv) of Table 7. The quadratic term is positive and statistically significant in the adoption equation, which cuts against the presence of an inverted-U relationship in the cement industry. The net effect of the competition regressors on adoption is negative throughout the support of the data, so the results should not be misinterpreted as implying that a sufficiently high degree of competition increases the benefits of competition. The analysis again provides corroborating support for the motivating theory.

6.3 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

\footnote{Another 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 perhaps competitors definitely would not adopt with low enough demand and definitely would adopt with high enough demand, which isolates intermediate ranges of demand as possibly supporting preemption. We implement be adding a quadratic in \textit{Construction}. The quadratic term has a t-statistic near zero in each of the baseline regressions, which again does not provide support for the importance of preemption in the data.}

32
<table>
<thead>
<tr>
<th></th>
<th>Adopt Precalculator vs. Maintain Old Kiln</th>
<th>Shut Down vs. Maintain Old Kiln</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i)</td>
<td>(ii)</td>
</tr>
<tr>
<td>Cost Savings</td>
<td>0.041***</td>
<td>0.041***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Fuel Costs</td>
<td>0.004</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Competitors</td>
<td>-0.039***</td>
<td>-0.073***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Precalculator Competitors</td>
<td>-0.012</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Competitors$^2$</td>
<td>0.0007***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>0.052***</td>
<td>0.046***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.011)</td>
</tr>
</tbody>
</table>

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 precalculator 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), there are two first stage regressions, for the number of competitors and the number of precalculator competitors, respectively. The excluded instruments are 20-year lags on the competition variables. In columns (ii), 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.
identical results. The linear probability model also returns similar effects in terms of both magnitude and statistical significance. We have also 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 again are consistent with the baseline multinomial probit model. Thus, the estimates appear to be driven by the empirical variation in the data rather than any 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 measures of the cost savings variable, based on fossil fuel prices five years ahead and behind the year of the observation. The alternative measures do not predict adoption. Column (ii) modifies the number of competitors based on a tighter distance radius of 200. This does not affect 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 $-0.015 - 0.041 = -0.056$. The results suggest that 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 much across these different time periods.

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 (e.g., the relevant mean elasticity falls from $-1.57$ to $-0.69$). 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. However, the 15-year lag produces results that are quite similar to the 20-year lag: the mean elasticities of adoption with respect to competition are $-1.41$ and $-1.57$, and the difference is not statistically significant. This result provides support for the 20-year lag as a valid instrument: any autocorrelation in the structural error terms appears to die out within 15 years. Lastly, the $F$-statistics in the baseline specifications range from 2,474 to 3,417 so the 20-year lag has power.
<table>
<thead>
<tr>
<th></th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
<th>(v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Savings</td>
<td>0.038***</td>
<td>0.015**</td>
<td>0.032***</td>
<td>0.030***</td>
<td>0.040**</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Cost Savings (t + 5)</td>
<td>-0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost Savings (t − 5)</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competitors (d &lt; 400)</td>
<td>-0.027***</td>
<td>-0.015**</td>
<td>-0.028**</td>
<td>-0.030**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Competitors (d &lt; 200)</td>
<td></td>
<td>-0.071***</td>
<td>-0.041***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.014)</td>
<td>(0.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>0.030***</td>
<td>0.085***</td>
<td>0.042***</td>
<td>0.041**</td>
<td>0.032**</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.017)</td>
<td>(0.009)</td>
<td>(0.018)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>IV Lag Structure</th>
<th>No IV</th>
<th>5 Year</th>
<th>10 Year</th>
<th>15 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Derived Statistics: Mean Elasticities of Pr(Adoption)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRT Cost Savings</td>
<td>0.52</td>
<td>0.54</td>
<td>0.55</td>
<td>0.56</td>
</tr>
<tr>
<td>WRT Competitors</td>
<td>-0.69</td>
<td>-0.92</td>
<td>-1.15</td>
<td>-1.37</td>
</tr>
<tr>
<td>WRT Construction</td>
<td>0.80</td>
<td>0.98</td>
<td>1.07</td>
<td>1.16</td>
</tr>
</tbody>
</table>

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.
7 Conclusion

The research discussed herein explores the determinants of technology adoption in the port-
land cement industry. The technology in question, the precalciner kiln, reduces the marginal
costs of production. The empirical results suggest that adoption is more likely if the cost sav-
ings are large, and less likely if there are many nearby competitors. The results also suggest
that competition damps the positive effect of cost savings on technology adoption. These
results are consistent with the predictions of a two-stage theoretical model in which firms
consider adoption, and then compete in accordance with Nash-Cournot competition. The
mechanism is simple: competition denies firms the scale required to recoup the investment
costs of technology adoption. The results thus support the relevance of the Schumpeterian
hypotheses regarding firm size and innovation, within a specific market environment.

We conclude with a brief discussion of two possible research extensions. First, our em-
pirical results suggest that the dynamic adjustment path of the cement industry in response
to market-based regulation of CO$_2$ likely would involve some combination of investment and
exit. Modeling this adjustment would be an interesting exercise. While some aspects of
this thought experiment have been studied (e.g., Fowlie, Reguant and Ryan (2016)) addi-
tional progress could be made. New results could have bearing on the welfare consequences
of regulation, and how market participants could most appropriately be compensated for
losses. Second, when evaluated together with other recent research, our results are sugges-
tive that preemption incentives are weaker in the presence of many competitors. However,
the conditions under which preemption speeds technology adoption remain under-explored
in the empirical literature. Further research could clarify the medium-run and long-run
relationships between competition and technology adoption.
References


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
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 Focal firm model**

What must be shown is that $\frac{\Delta b_i(\Delta c, c_i^0, \bar{c}, a, N)}{\Delta N} < 0$ under condition (5), because it is evident from equation (4) that increasing $N$ reduces the (positive) derivative of benefits with respect to $\Delta c$ under condition (5). Benefits decrease in $N$ if

$$b_i(\Delta c, c_i^0, \bar{c}, a, N + 1) < b_i(\Delta c, c_i^0, \bar{c}, a, N)$$

Plugging in equations (3) and (4) yields

$$\frac{N + 1}{N + 2} a - \hat{c}_i + (N + 1)(\bar{c} - \hat{c}_i) < \frac{N}{N + 1} a - \hat{c}_i + N(\bar{c} - \hat{c}_i)$$

and with some effort this can be manipulated to obtain condition (5).
B.2 Long run model

Derivation of equation \(6\). We solve for the SPE number of adopters, \(L^*\), which is characterized by the inequalities \(k \leq q_1(L = L^*)^2 - q_0(L^* - 1)^2\) and \(k > q_1(L^* + 1)^2 - q_0(L^*)^2\). (We have condensed the notation for brevity.) We start with the first inequality. It can be reexpressed:

\[
k \leq (q_1(L^*) + q_0(L^* - 1)) \times (q_1(L^*) - q_0(L^* - 1))
\]

because \(x^2 - y^2 = (x + y)(x - y)\). It can be shown than \(q_1(L) = q_0(L) + \Delta c\) for any \(L\). Thus, plugging in for \(q_1\) yields

\[
k \leq (q_0(L^*) + \Delta c + q_0(L^* - 1)) \times (q_0(L^*) + \Delta c - q_0(L^* - 1))
\]

If there are \(L\) firms with costs of \(c^1\) and \(N - L\) firms with costs of \(c^0\), then equation \(3\) can be used to obtain the equilibrium quantity of a non-adopter:

\[
q_0(L) = \frac{a - N c^0 + L c^1 + (N - L - 1) c^0}{N + 1} = \frac{a - c^0 - L \Delta c}{N + 1}
\]

Substituting into the inequality yields

\[
k \leq \left(2 \frac{a - c^0}{N + 1} - \left(\frac{L}{N + 1} + \frac{L - 1}{N + 1}\right) \Delta c + \Delta c\right) \left(- \frac{L}{N + 1} \Delta c + \Delta c + \frac{L - 1}{N + 1} \Delta c\right)
\]

Collecting terms,

\[
k \leq \left(2 \frac{a - c^0}{N + 1} + \frac{N - 2L}{N + 1} \Delta c\right) \left(\frac{N}{N + 1} \Delta c\right)
\]

The second inequality that characterizes \(L^*\) is analogous, so we also have

\[
k > \left(2 \frac{a - c^0}{N + 1} + \frac{N - 2(L + 1)}{N + 1} \Delta c\right) \left(\frac{N}{N + 1} \Delta c\right)
\]

Combining the inequalities yields

\[
\frac{N - 2(L + 1)}{N + 1} < \frac{k}{(\Delta c)^2} \frac{N + 1}{N} - \left(2 \frac{a - c^0}{\Delta c}\right) \left(\frac{1}{N + 1}\right) \leq \frac{N - 2L}{N + 1}
\]

Multiplying by \(N + 1\) and subtracting \(N\) yields

\[
-2(L + 1) < \frac{k}{(\Delta c)^2} \frac{(N + 1)^2}{N} - 2 \frac{a - c^0}{\Delta c} - N \leq -2L
\]
Finally, dividing by negative two flips the direction of the inequalities

\[ L \leq N + \frac{a - c_0}{\Delta c} - \frac{1}{2} \frac{k}{(\Delta c)^2} \frac{(N + 1)^2}{N} < L + 1 \]  

(B.5)

This is equation (9) in the text.

**Proof of Result 3.** If \( k > 2(\Delta c)^2 \), it can be shown that as long as not everyone adopts, then the number of adopters decreases with competition in some interval. To see this, we work with middle part of the inequality in equation (9) and take the derivative of the bound

\[ N + \frac{a - c_0}{\Delta c} - \frac{1}{2} \frac{k}{(\Delta c)^2} \frac{(N + 1)^2}{N}, \]

which is

\[ 1 - \frac{1}{2} \frac{k}{(\Delta c)^2} \frac{N^2 - 1}{N^2}. \]

We want to know when this derivative is negative, equivalently, finding the zeros of this equation. Note that the zeros of this equation solve

\[ N^2 \left( 1 - \frac{1}{2} \frac{k}{(\Delta c)^2} \right) + \frac{1}{2} \frac{k}{(\Delta c)^2} = 0. \]

The above equation defines a parabola in \( N \) that will slope upwards if \( k < 2(\Delta c)^2 \) and downwards if \( k > 2(\Delta c)^2 \). The parabola is maximized or minimized at \( N = 0 \). The roots of the equation are

\[ \pm \sqrt{\frac{-\frac{1}{2} \frac{k}{(\Delta c)^2}}{1 - \frac{1}{2} \frac{k}{(\Delta c)^2}}} \]

If \( k < 2(\Delta c)^2 \) then this equation has no real roots. Thus the parabola slopes upwards, is always positive, and more competition always increases the incentive to adopt. If instead it slopes downwards, there is one real positive root which I’ll call \( N_1 \). If \( N > N_1 \) and \( L^* + 1 \leq N \), then it must be that increasing \( N \) weakly decreases the number of firms in the market (the inequality can’t be strong due to the integer nature of the number of firms). Note that there will be some point at which the upper bound on \( L^* \) becomes negative, since the upper bound is decreasing in \( N \), and the upper bound on \( L^*/N \) converges to a negative number as \( N \) goes to infinity. This means there is some number \( N_2 \) such that for \( N > N_2 \) nobody adopts. Within the interval \( N_1 \) and \( N_2 \), since \( L \) decreases as \( N \) rises, the fraction of adopters must also decrease. The bound \( N_1 \) is unnecessarily tight, as the fraction of firms
that adopt could decrease in \( N \) even if \( L \) is increasing in \( N \) – it is sufficient that \( L \) increases in \( N \) at a slower rate than \( N \).

### B.3 Dynamic model

This section shows the intermediate steps needed to derive the approximation to the dynamic benefit of upgrading in equation (15) of Section 2.3. We start by noting that we can solve for the future value of upgrading, \( V_1(N) \), explicitly, since once a firm has upgraded the decision is irreversible. In particular, the value function vector can be written as

\[
V_1(N) = \Pi_0 \left( \pi_1 + \delta \left( \sum_{t=0}^{\infty} \delta^t \Pi_1^t \pi_1 \right) \right)
\]

\[
= \Pi_0 (I + \delta (I - \delta \Pi_1)^{-1}) \pi_1,
\]

It is slightly more complicated to solve for \( V_0(N) \), however, we can apply a similar idea to the one above. To get started note that element \( L \) of \( V_0(N) \) will be

\[
V_{0,L}(N) = F(k^*(L))v_1(L; N) + (1 - F(k^*(L)))v_0(L; N)
\]

\[
= P_0(L)'(F(k^*(L))((I + \delta \Pi_0 (I + \delta (I - \delta \Pi_1)^{-1})))\pi_1 +
\]

\[
+ (1 - F(k^*(L))((\pi_0 + \delta V_0(N)))) - F(k^*(L))E(k|k < k^*(L))
\]

\[
= P_0(L)'(F(k^*(L))((I + \delta \Pi_0 (I + \delta (I - \delta \Pi_1)^{-1})))\pi_1 +
\]

\[
+ (1 - F(k^*(L))((\pi_0 + \delta V_0(N)))) - F(k^*(L))E(k|k < k^*(L))
\]

Note that if we write the left hand side of the equation above it defines an equation for \( V_0(N) \):

\[
V_0(N) = \Pi_0 \left( F(I + \delta \Pi_0 (I + \delta (I - \delta \Pi_1)^{-1})) \right) \pi_1 + (I - F)\pi_0 + \delta(I - F)V_0 + F\kappa,
\]

where we define \( F \) to be a diagonal matrix with \( F(k^*(L)) \) on each \( L + 1 \)st diagonal element, and \( \kappa \) is a diagonal matrix with \( E(k|k < k^*(L)) \) on each \( L + 1 \)st diagonal element. If we solve this equation for \( V_0 \) then we can express \( V_0 \) as

\[
V_0 = A\pi_1 + B\pi_0 + C,
\]
where

\[ A = (I - \delta \Pi_0 (I - F))^{-1} \Pi_0 F (I + \delta \Pi_0 (I + \delta (I - \delta \Pi_1)^{-1})) \]
\[ B = (I - \delta \Pi_0 (I - F))^{-1} (I - F) \]
\[ C = (I - \delta \Pi_0 (I - F))^{-1} F \kappa \]

We can use the above information to compute the benefit of upgrading at any state, \( b \), as

\[ b = \Pi_0 (\pi_1 + \delta V_1 - \pi_0 - \delta V_0) \]
\[ = \Pi_0 ((I + \delta (\Pi_0 (I + \delta (I - \delta \Pi_1)^{-1})) \pi_1 - \pi_0 - \delta A \pi_1 - \delta B \pi_0 - \delta C) \]

We can plug into the equation above the Taylor approximation for \( \pi_1 \), which is

\[ \pi_1 \approx \pi_0 + \frac{\partial \pi}{\partial s} \Delta c, \]

and after a bit of algebra one can derive equation (15).

C Additional Figures and Tables
Table C.1: Probit Regression Results with Interaction

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Equation for Shut Down vs. Maintain</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fuel Costs, Competition, and Demand</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuel Costs</td>
<td>0.009</td>
<td>0.006</td>
<td>-0.022*</td>
<td>-0.032***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.013)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>Competitors</td>
<td>0.008</td>
<td>0.004</td>
<td>-0.015</td>
<td>-0.026**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>Fuel Costs</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001***</td>
<td>0.001***</td>
<td></td>
</tr>
<tr>
<td>\times Competitors</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>-0.027**</td>
<td>-0.026**</td>
<td>-0.018</td>
<td>-0.005</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.009)</td>
<td></td>
</tr>
</tbody>
</table>

**Mean Elasticities of Pr(Shut Down)**

| 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  |

**Specification Details**

| 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.