Instability of Endogenous Price Dispersion Equilibria: a Simulation

Lucas Herrenbrueck
Simon Fraser University

First WP version: 2015. This version: 2017
Accepted at the Canadian Journal of Economics

Abstract

Models of price posting by firms and search by consumers (such as Burdett and Judd, 1983), often feature equilibria with endogenous price dispersion. However, such equilibria are strategically fragile. In order to investigate how robust they are in the absence of an external coordination mechanism, I simulate various protocols firms may use to update their prices. Despite firms being myopic, some protocols yield results close to the benchmark model. If firms rush to update before observing competitors' actions, profits are higher on average but volatile and cyclical. With cost dispersion, prices become more stable as they are more closely tied to costs. All results are robust to moderate menu costs.

JEL codes: D21, D43, D83

Keywords: Simulation; search frictions; Burdett-Judd pricing; menu costs; disequilibrium dynamics

Contact: herrenbrueck@sfu.ca, +1-778-782-4805
Simon Fraser University, Department of Economics
8888 University Drive, Burnaby, B.C. V5A 1S6, Canada

I am grateful to Paul Bergin, Giacomo Bonanno, Athanasios Geromichalos, Martine Quinzii, Katheryn Russ, Burkhard Schipper, Joaquim Silvestre, and Ina Simonovska, as well as several anonymous referees, for their helpful comments and suggestions.

1 Introduction

Economic models almost universally focus on equilibrium. This focus is justified to the extent that at least over a sufficiently long time horizon, beliefs converge to facts and best response opportunities have been consummated. But partly, it reflects the limitations of mathematical modeling. Disequilibrium is hard to define and harder to compute. The stability of equilibrium, how disequilibrium converges to equilibrium, and the properties of long-run and large-scale aggregates are therefore important questions of economic theory.

In this paper, I consider a standard model of endogenous price dispersion (Burdett and Judd, 1983; Mortensen, 2005) and simulate it with a finite number of firms – and, crucially, without any external coordination mechanism. The firms are identical and compete for the sale of a homogeneous good, like in a Bertrand model, which would tend to drive prices down to marginal costs and profits to zero. But consumers face search frictions: at a given time, they will only observe a finite menu of prices, which means that firms can make positive profits by charging a very high price and selling only to consumers who do not have a better option as part of their menu. This tension implies that price dispersion is present in any equilibrium, and the distribution of prices will adjust endogenously to make firms indifferent between charging high prices for low sales, and charging low prices for high sales.

With a finite number of firms, it is impossible to make firms exactly indifferent. They would like to charge prices just below but as close as possible to their competitors, which, depending on the exact form of the strategy space, could lead to repeated cycles of under-cutting followed by large price rises (Edgeworth cycles). This will make the prices of individual firms volatile, and possibly aggregates like the price distribution, and the level and dispersion of profits, too.

I therefore simulate various strategic protocols firms may use to update their prices to answer the following questions: (1) Starting from a random price distribution, is there "tâtonnement", i.e. convergence to any equilibrium? (2) The benchmark model of a continuum of firms predicts an equilibrium where firms make equal and constant profits. Does the simulation converge to this equilibrium? (3) How do the answers depend on the number of firms in the simulation? (4) How do the answers depending on a richer structure of costs, including menu costs?

The results can be summarized as follows: convergence happens, but only to an extent. The average level of profits is higher than in the benchmark model even when very many firms are simulated, this level is volatile over time, and there is substantial profit dispersion. In some simulations, the results are well-behaved: first, the volatility and dispersion of profits are in line with what we would see if firms were playing a mixed strategy equilibrium; second, the upward bias of prices and profits decreases as the number of firms gets large. In an alternative specification in which firms are extremely myopic, the divergence from the benchmark equilibrium is stark: profits are excessively high, dispersed, volatile, and cyclical, and this does not change as the number of firms gets large.

The idea that every firm would like to undercut its competitors by the smallest possible margin, the "Bertrand paradox", has been studied in a literature too large to survey here. A relevant example is Maskin and Tirole (1988), who demonstrate that Edgeworth cycles can arise as equilibria of dynamic Bertrand games with a small number of firms, supported by a folk theorem. A key difference in the approach of Burdett and Judd (1983) and Burdett and Mortensen (1998) is that marginal cost pricing is not an equilibrium even with infinitely many firms, because profits can be made by charging a very high price and still attracting a positive measure of consumers.[1]

The oldest approach to understanding the stability and convergence properties of economic equilibrium is Walrasian tâtonnement.[2] There has also been some work on models with dispersed price equilibria. Hopkins and Seymour (2002) approach a similar problem theoretically and find that the dispersed equilibrium can be stable under learning if competition is not too severe. In several simulations and experiments, however, Cason, Friedman, and Wagener (2005) found that firms' actions did not converge to a dispersed price equilibrium, but exhibited Edgeworth cycles, and the amplitude of these cycles did not tend to decrease over time.

And understanding the stability properties of endogenous price dispersion is important, because it is more than a curiosity from economic theory. The Burdett-Judd model, in particular, has been successfully applied in many contexts with wide relevance. For example, Head and Kumar (2005) and Head, Kumar, and Lapham

[1] This approach is often termed *random search* because the probability that a buyer meets (or more appropriately in this context, observes) a particular seller is independent of seller characteristics, such as price, capacity, or queue length. Using the main alternative paradigm, *directed search*, Geromichalos (2014) studies the Bertrand paradox and a general class of resolutions of the paradox.

[2] A good summary of the theory of Walrasian tâtonnement is provided in Chapters 17 and 20 of Mas-Colell, Whinston, and Green (1996). In recent work, Crockett, Oprea, and Plott (2011) provide experimental support of the theory.

(2010) looked at cyclical variation in markups; Alessandria (2009) and Herrenbrueck (2015) modeled price dispersion in international trade; Head, Liu, Menzio, and Wright (2012) and Burdett and Menzio (2013) showed that Burdett-Judd price dispersion can explain the pattern of price changes in the macroeconomy without resorting to ad-hoc stickiness; Kaplan and Menzio (2016) showed that differences in search behavior between employed and unemployed workers contribute to aggregate fluctuations; Menzio and Trachter (2015) and Kaplan, Menzio, Rudanko, and Trachter (2016) studied price dispersion in retail; and Burdett, Trejos, and Wright (2015) argued that monetary-search models with price posting rather than bargaining are both more realistic and more tractable. Since my simulations show that the Burdett-Judd benchmark is robust under some strategic protocols but not under others, and that efforts by uninformed firms to grope their way towards strategic improvements can themselves contribute to aggregate fluctuations, the paper has relevance for all of this literature.

2 Benchmark model

The model is a partial equilibrium version of Burdett and Judd (1983). The number of firms is modeled either as finite $N < \infty$, or as a continuum in which case the set of firms has unit measure. The measure of consumers will be described later. Consumers value a good that each firm can produce at constant marginal cost c. Consumers have one unit of money, and they will spend all of it unless the price exceeds a reservation price $\bar{p}$.[3] Meetings between consumers and firms are subject to search frictions: firms post prices before meetings take place, consumers observe a random number of prices ("receive quotes") and spend all of their money on the firm with the lowest price, unless that price exceeds the reservation price.

The random number of price quotes, K, can be described by the probability mass function $q_k = Pr\{K = k\}$ with support $\{0, \ldots, \bar{k}\}$. For example, Burdett and Judd (1983), Head et al. (2012), and Wang (2016) assume that consumers receive either 1 or 2 quotes. Others, following Mortensen (2005), assume that K follows a Poisson distribution; Burdett et al. (2015) also use a Logarithmic distribution. For reasons of generality, I prefer the flexible form where K follows a negative binomial distribution with parameters $1/\rho$ and $\eta\rho/(1 + \eta\rho)$. This means that the expected number

[3] This demand curve arises optimally in Head and Kumar (2005). Here, I use it for simplicity.

of quotes is η, and higher ρ increases the dispersion of both quoted and transaction prices, which nests all of these cases.[4] For example, the limit $\rho \to 0$ corresponds to Poisson, the limit $\rho \to \infty$ (after rescaling η) corresponds to the Logarithmic case, and $\rho = -0.5$ with $\bar{k} = 2$ corresponds to the original Burdett-Judd model.

In order to normalize the measure of matched consumers to 1, assume that the overall measure of consumers is $(1 - q_0)^{-1}$.

First, consider a continuum of firms, and let the CDF of posted prices be $F(p)$. No firm would find it optimal to charge more than the reservation price in, which implies $F(\bar{p}) = 1$ in equilibrium. Therefore, every consumer will make a purchase, and will do so from the firm with the lowest price. The CDF of transaction prices will be $J(F(p))$, where J is the concave function:

$$J(F) = \sum_{k=0}^{\infty} q_k \left[1 - (1 - F)^k\right] = \frac{1 - (1 + \rho\eta F)^{-1/\rho}}{1 - (1 + \rho\eta)^{-1/\rho}} \tag{1}$$

If F has no mass points, a firm charging price $p \leq \bar{p}$ can expect to sell to $J'(F(p))$ consumers. Recall that each customer has unit elastic demand so that the *expenditure* is fixed, as long as the price does not exceed the reservation price. Therefore, the profits of a firm with price $p \leq \bar{p}$ satisfy:

$$\pi(p) = \left(1 - \frac{c}{p}\right) J'(F(p)) \tag{2}$$

Burdett and Judd (1983) prove that endogenous price dispersion is the only equilibrium as long as both $q_1 > 0$ and $q_2 > 0$, and that some firms do charge the reservation price ($F(p) < 1$ for $p < \bar{p}$). But because all firms are the same ex-ante, they must be indifferent between charging any price on the support of F; in particular, they must be indifferent to charging $\bar{p}$ and making the minimal sales $J'(1)$:

$$\pi_\infty = \left(1 - \frac{c}{\bar{p}}\right) J'(1) = \left(1 - \frac{c}{p}\right) J'(F(p)) \tag{3}$$

(The index ∞ stands for a continuum of firms.) Using this equation, we can solve for the (partial) equilibrium price distribution, $F_\infty(p)$, taking c and $\bar{p}$ as given:

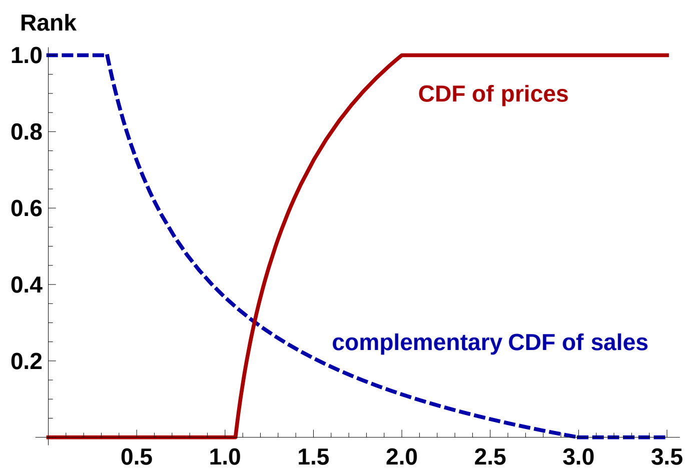
[4] The parameters need to satisfy $\eta > 0$ and $\rho > -1/\eta$. If $\rho \geq 0$, $\bar{k} = \infty$, otherwise $\bar{k} = -1/\rho$. This class of matching processes is derived in detail in Herrenbrueck (2015).

$$F_\infty(p) = \frac{1+\rho\eta}{\rho\eta}\left[\frac{1-c/p}{1-c/\bar{p}}\right]^{\frac{\rho}{1+\rho}} - \frac{1}{\rho\eta} \tag{4}$$

for $p \in [\underline{p}, \bar{p}]$, where $\underline{p} = \arg\inf_p \{F_\infty(p) > 0\}$.

Figure 1: **The endogenous price and sales distributions**

For firms of any rank, profits (sales multiplied by the profit margin) are constant.



If the number of firms is finite, then there exists no pure strategy Nash equilibrium, but there still exist mixed-strategy equilibria. Since the number of price quotes k that a consumer sees is now sampled from finitely many firms, however, we have to ask if this sampling is with or without replacement. Both of these are legitimate ways of "discretizing" the Burdett-Judd model, both have the same price and profit distribution in the $N \to \infty$ limit, but they have different properties for finite N.

Consider first sampling without replacement. Then a consumer with $k = N$ must see *all* price quotes; and $k > N$ is impossible so we cannot directly use the unbounded distribution of quotes described above. We could truncate this distribution at $k \leq N$, but this implies that having more firms relaxes the truncation mechanically.[5] Furthermore, profits as a function of the individual firm's price become intractable for large N.

[5] The ultimate implication is unclear, because it would depend on whether q_1 would be scaled up or down relative to all q_k, $k \geq 2$. And conditional on a *fixed* quote distribution $\{q_k\}_{k=0}^{\bar{k}}$, the equilibrium price distribution and expected profits are equal to F_∞ and π_∞ – thus, independent of N – when quotes are sampled without replacement.

Alternatively, we can consider sampling with replacement. To illustrate, consider $N = 2$ firms, and suppose $q_1 = q_2 = 0.5$, all other $q_k = 0$. What are the chances a consumer who samples two prices $k = 2$ sees *both* prices? Only 50%. What are the chances that, despite sampling two prices, a consumer is stuck with only a quote of the higher price of the two firms? It is 25%, not 0%, because sampling is with replacement. For a simple analogy, suppose prices are advertised in TV commercials: even a consumer who watches ten commercials has a small chance of seeing the same one ten times. In this case, the equilibrium price distribution becomes intractable for large N, but the profit function is easy to characterize (De los Santos, Hortaçsu, and Wildenbeest, 2012). Let $R(p) \in \{1, \ldots, N\}$ be the realized rank of price p among all prices. Then the realized profits of a firm charging p are:

$$\pi(p) = \left(1 - \frac{c}{p}\right) \cdot \left[J\left(\frac{R(p)}{N}\right) - J\left(\frac{R(p) - 1}{N}\right)\right] \tag{5}$$

And in a mixed strategy equilibrium, all firms must still be indifferent to charging the reservation price, $\bar{p}$. At that price, a firm is guaranteed to have price rank N; therefore, the benchmark level of total profits with N firms, multiplying the equilibrium profit level by the number of firms, is:

$$\pi_N = \left(1 - \frac{c}{\bar{p}}\right) \cdot \left[J\Big(1\Big) - J\left(1 - \frac{1}{N}\right)\right] \cdot N \tag{6}$$

Taking the limit as $N \to \infty$, we see that the terms in (5, 6) indeed converge to (2, 3).

With a finite number of firms drawing prices randomly from an equilibrium distribution, we must expect some volatility and dispersion of realized profits. Because the price distribution with N firms is not tractable, I will use an approximate simulation of random draws from the limit price distribution $F_\infty(p)$ to obtain a handle on how much noise we should expect from mixing alone.

Tractability is not a problem for $N = 2$, where we can obtain exact solutions whether sampling is with replacement or without. For additional intuition, this case is solved explicitly in Appendix A.1 and compared to simulation results.

3 Simulation models

There are N firms. Each simulation consists of a random initial state, representing the firms' prices, and an updating procedure to map today's state into tomorrow's. I consider several different updating procedures to allow for flexibility regarding how firms change their prices and what they know. Sales and profits are computed using Equation 5. Simulation time is discrete.

Non-atomic firms are an essential ingredient of the Burdett-Judd model, because only a non-atomic price distribution can support equilibrium (corner cases where q_1 equals 0 or 1 aside). The reason is that the *distance* between observed prices does not affect consumer choice: consumers make all their purchases from the cheapest firm that they observe. Consequently, any firm would like to price as close to the next-most expensive firm as possible, but strictly below. However, because not all consumers observe more than one price, there are always positive profits to be made by charging the reservation price. There exists no pure-strategy Nash equilibrium with a discrete number of firms.

This matters because mixed-strategy equilibria are not easy to find for simple computational decision makers, or indeed complex ones (like humans). In the Burdett-Judd environment, this difficulty is amplified: the best response function of a firm is indeterminate in equilibrium (because firms must be indifferent), and not at all continuous in other firms' actions, in equilibrium or out of it. (For example, suppose all firms charging less than $\bar{p} - \varepsilon$ shifted their price up by ε. Then the best response of firms $[\bar{p} - \varepsilon, \bar{p}]$ would be to shift all the way down to the bottom of the distribution.) So it becomes an interesting question whether firms following a simple profit improvement algorithm can "find" the Burdett-Judd benchmark equilibrium, or something close to it – and if not, what they might "find" instead.

With this goal, I investigate a variety of simulation models. The first model uses the following approach: every simulation period, a single firm learns what its profits would be if it charged a particular price randomly chosen from a continuum. If these profits exceed the current profits, it switches to the new price, otherwise it keeps the current price. The other models offer variations on this theme. In detail:

Model 1

The set of possible prices is the interval $[c, \bar{p}]$. Firms begin with prices randomly selected from the interval. In random sequential order, they draw one potential price from the same interval. They compute the rank of both their current price and of the potential new price (defined as N for the cheapest price and 1 for the most expensive), and compare profits according to the formula in Equation (5). Their price rank is well-defined because almost-surely no two firms charge the same price. They switch to the new price if and only if profits would be higher than at the current price.[6]

The goal of Model 1 is to introduce a computational approximation of the Burdett-Judd environment, with strategic frictions imposed on decision makers: they can only sample one potential improvement at a time, they are myopic, they are restricted to pure strategies. The question is how closely this simulation is able to mimic the Burdett-Judd benchmark, and how this depends on the number of firms, N.

Model 2

The environment is the same as in Model 1, but in each time period, all firms make their choices <u>at once</u>, and in a very unstructured manner: they take into account the effects on the price distribution of their own potential switch, but not the fact that other firms may switch, too. (It would be possible to design a procedure where only a few firms switch at a time without taking each other's choices into account, but this wholesale switching procedure provides an extreme point of comparison.)

The goal of Model 2 is to make "rational" actions even harder compared to Model 1, as firms are required to update simultaneously as opposed to in a staggered order. The question is the same as for Model 1.

Model 3

At the beginning of time, each firm is permanently assigned a cost parameter from the interval $[c, \bar{c}]$, where $\bar{c} < \bar{p}$. The procedure for updating prices is the same as in Model 1.

The first goal of Model 3 is to see how critical the indifference property of Burdett-Judd competition with equal costs is to the simulation results. In the Burdett-Judd model with cost dispersion, each firm's equilibrium price is pinned down uniquely, but the equilibrium price distribution is still an endogenous object. The second goal is to see how closely the ranking of prices

[6] This choice is myopic because the updating firm does not anticipate that other firms will be changing their prices, too, before its next opportunity to update.

tracks the ranking of costs in the simulation; in the benchmark, the rankings should be identical.

Model 4

Firms sample possible prices as in Models 1-3; however, they face a menu cost $d \in (0, \pi_\infty)$, and will only update their price if the difference between new and old profits exceeds this menu cost.

The goal of Model 4 is to compare how frequently firms update their prices as a function of the menu cost, and how menu costs affect the aggregate outcomes of the simulations.

Models A.1 and A.2

These models are described and analyzed in Appendix A.2. As an alternative to seeing a single, randomly sampled, price, firms choose prices freely from a discrete grid.

The reservation price is set to $\bar{p} = 2$, and the matching process is parametrized by mean $\eta = 2$ and shape $\rho = 1$. Models 1 and 2 are simulated for $N = 16$, $N = 40$, and $N = 100$, with cost $c = 1$ for all firms. (Model 1 is also simulated for $N = 6$ and $N = 625$ to provide extreme points of comparison.) Model 3 is simulated for $N = 16$, with costs equally spaced between $c = 1$ and one of $\bar{c} = 1.1$, $\bar{c} = 1.4$, and $\bar{c} = 1.9$. Finally, Model 4 compares the $N = 16$ simulations of models 1, 2, and 3 (the middle case with $\bar{c} = 1.4$), for menu costs of $d = 0.1\pi_\infty$, $d = 0.2\pi_\infty$, and $d = 0.4\pi_\infty$.[7] Each model is simulated for 21,000 periods, and after discarding the first 1000 to make sure the results do not depend on the random initial state, each reported simulation consists of 20,000 periods.

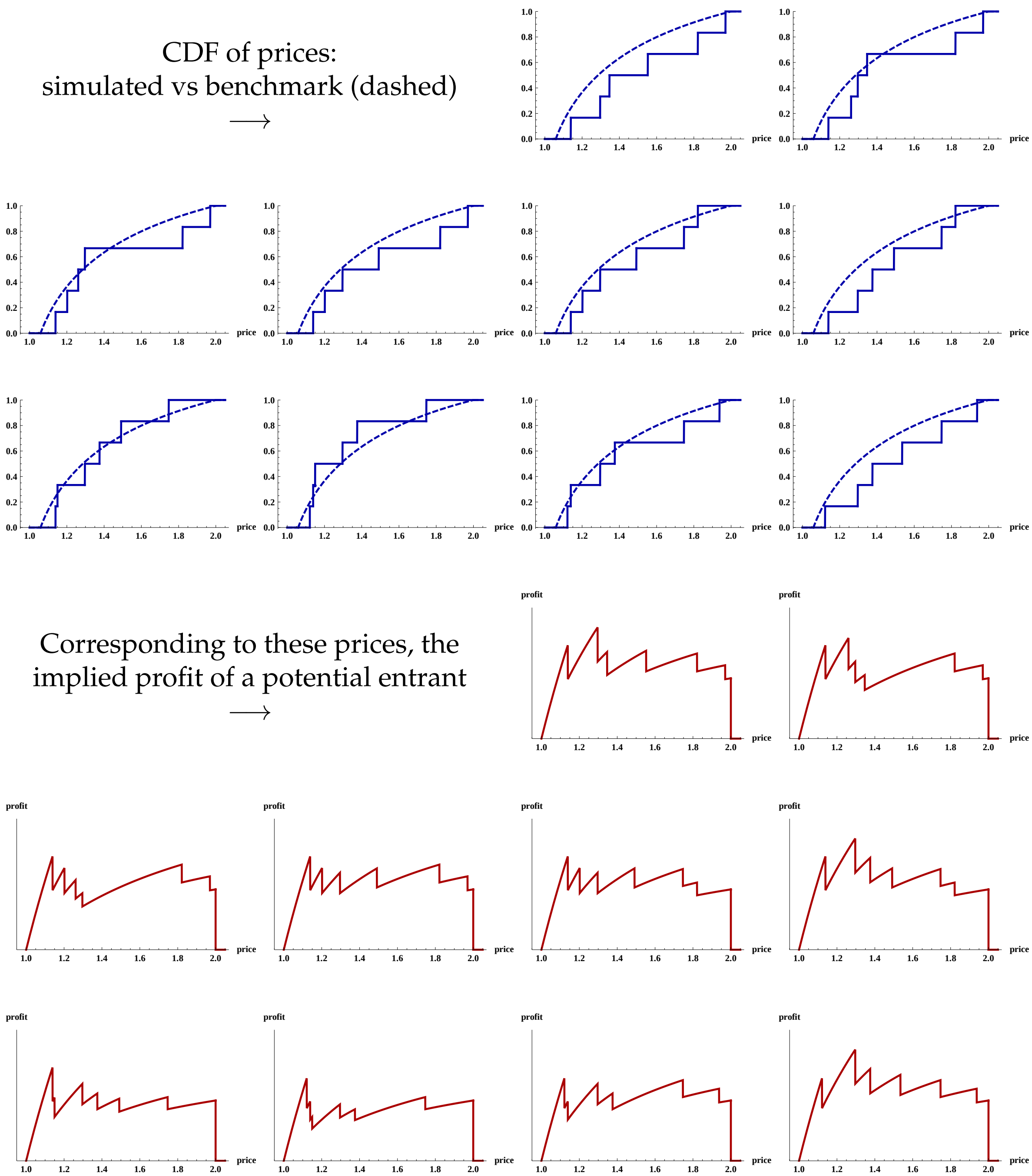
With these parameters, Equation (5) yields the following benchmarks for profits:

Table 1: **Profit benchmarks**

N	6	16	40	100	∞
π_N	0.187	0.174	0.169	0.168	0.167

[7] With 16 firms and equal marginal costs, realized profits typically vary between π_∞ and $1.5\pi_\infty$, which provides a point of comparison for the menu costs. Firms with higher marginal costs have lower profits, so they will be relatively more vulnerable to the menu cost; consequently, the menu cost parameter is adjusted downward in that particular case, to be a proportion of the *lowest* profits in the benchmark equilibrium.

Figure 2: **Excerpt from the simulation**



Excerpt from the simulation of Model 1 with $N = 6$, showing a particular sequence of 10 consecutive price updates. Simulation steps where a firm chose not to update are omitted.

4 Simulation results

Results from Model 1

Figure 2 shows a particular excerpt from the simulation of Model 1 with $N = 6$ to illustrate the principle. We can see that the prices of individual firms never converge because there are always switching opportunities, no matter how infinitesimal. The cumulative distribution functions (CDFs) of prices do not converge to a limit function, either, as there is always churn in the price ranking.

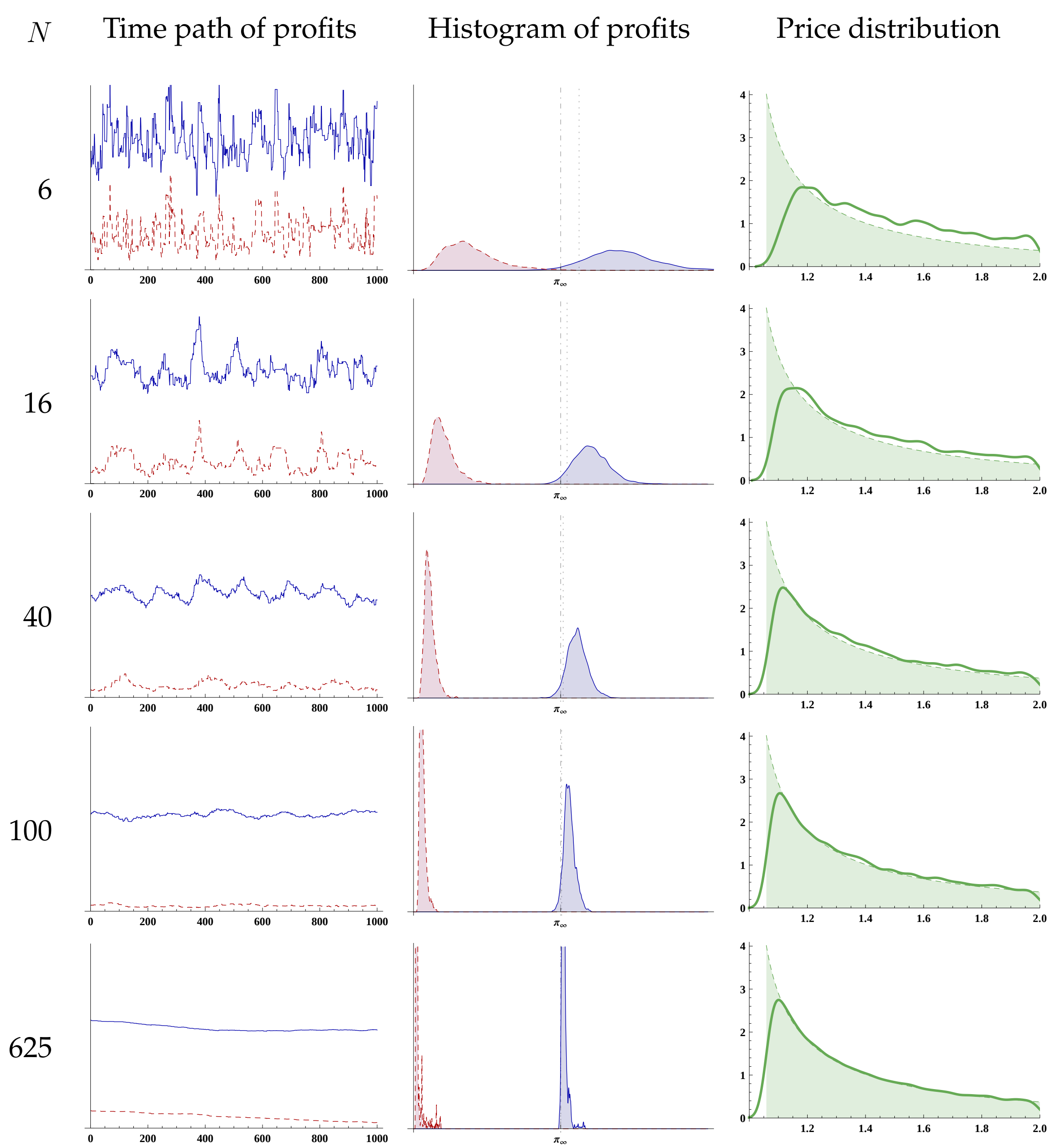
However, some churn must be expected, and the first question is *how much*. The simplest way to compute this (approximately) is to suppose all firms draw prices randomly from the limiting distribution $F_\infty(p)$ every period (since the counterpart $F_N(p)$ is not tractable), and simulate the resulting price and profit aggregates. These results are included in Table 2 below.

For the simulations of Models 1 and 2, where firms update prices subject to severe strategic frictions, it can still be the case that the benchmark model provided an accurate representation of prices and profits *over time*, even if they differ in any single period. If that is the case, then the long-run price distribution should converge to $F_\infty(p)$ as $N \to \infty$. Total profits should be approximately equal to π_N for $N < \infty$, they should converge to π_∞, and they should not vary too much over time. Cross-sectional dispersion of profits should not be too large, and vanish as $N \to \infty$.

Detailed results are reported in Table 2 and illustrated in Figure 3. Given the built-in strategic limitations (myopia, no memory, no foresight), the long-run averages of Model 1 are surprisingly close to the benchmark(s). Simulated profits are higher on average than in the benchmark model, and exhibit substantial volatility and dispersion. But volatility and dispersion are similar to what we obtain from random mixing alone, and all of these statistics fall as the number of firms becomes large. The average price distribution is robustly shifted up compared to the continuum benchmark, but it clearly converges as the number of firms gets large.

However, even with 100 firms, average profits remain at least 5% above any of the benchmarks, suggesting that the Burdett-Judd equilibrium is limited as a description of small-market competition under strategic frictions, even averaged over time. Another curious observation is that the correlation over (simulated) time between the cross-sectional mean and coefficient of variation of profits is robustly positive, which suggests that variation in profits is driven by occasional outliers to the

Figure 3: **Order (profits and prices in Model 1)**



Cross-section mean and standard deviation (dashed) of profits, truncated to the first 1000 simulation periods.

Smoothed histograms of the cross-section mean (right) and standard deviation (dashed, left) of profits. Vertical lines indicate profit benchmarks π_∞ and π_N.

Density functions of prices pooled over all simulation periods, compared with the benchmark equilibrium (shaded area).

Table 2: **Profits in Models 1, 2, and 5**

Model	N	Level	Volatility	Dispersion	L-D correlation
1	6	0.236	0.039	0.267	0.56
	16	0.201	0.020	0.167	0.46
	40	0.186	0.011	0.099	0.39
	100	0.177	0.006	0.057	0.43
	625	0.171	0.003	0.026	0.79
2	16	0.214	0.035	0.229	0.63
	40	0.206	0.039	0.205	0.72
	100	0.205	0.032	0.201	0.81
Random draws from $F_\infty(p)$	16	0.171	0.027	0.149	0.21
	40	0.168	0.016	0.096	0.13
	100	0.167	0.010	0.061	0.09
Continuum benchmark	∞	0.167	0.000	0.000	–

"Level" is the cross-sectional total of firms' profits in a simulated period, averaged across simulated time. "Volatility" is the time-series standard deviation of profit levels. "Dispersion" is the cross-sectional coefficient of variation of firms' profits in a simulated period, averaged across simulated time. The "L-D correlation" is the time-series correlation between the level of profits and their dispersion (measured by the coefficient of variation to avoid a mechanical correlation between means and standard deviations) in each simulated period.

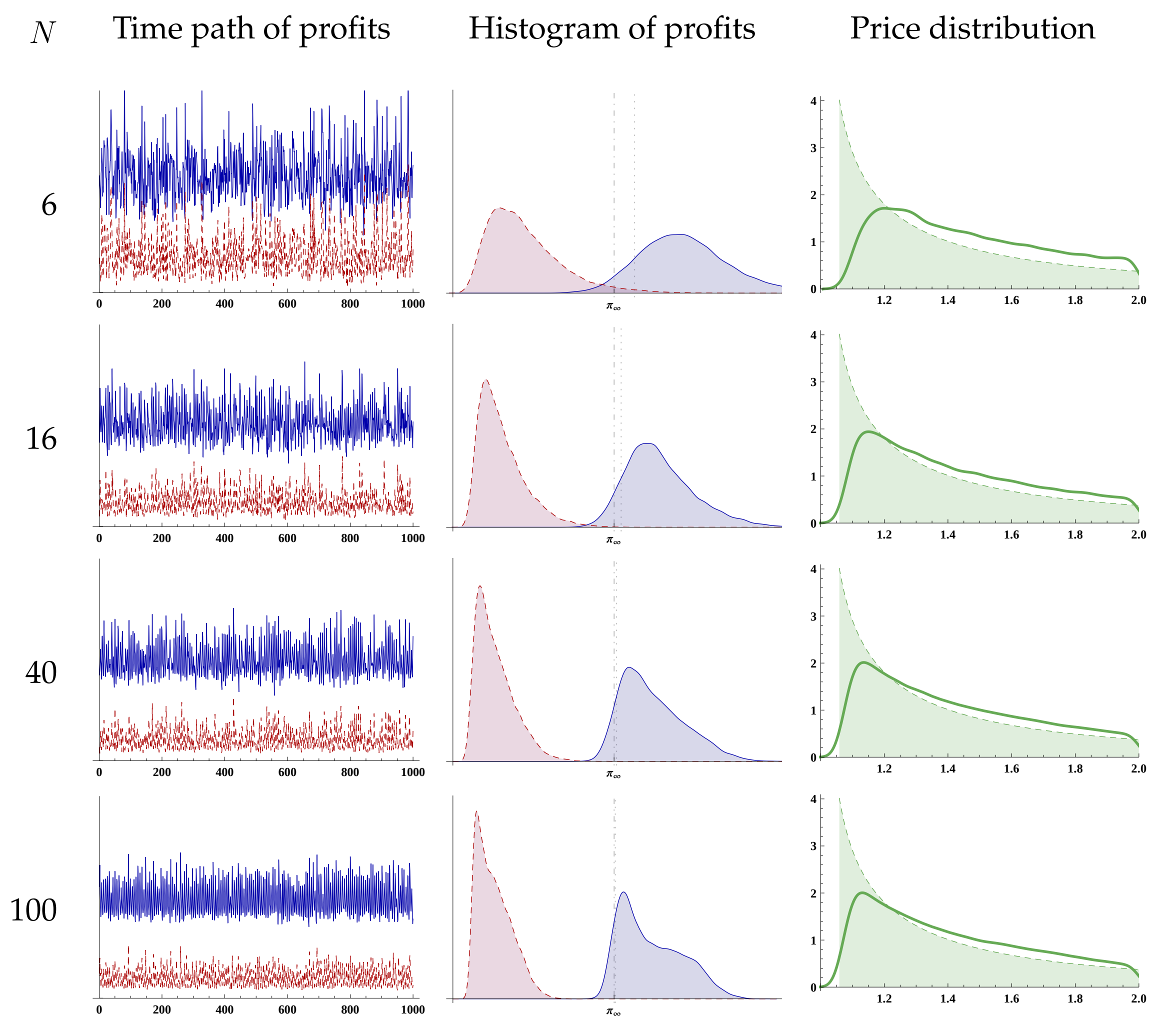
top and not to the bottom.

Results from Model 2

The simulation of Model 2 exhibits stranger behavior, illustrated in Figure 4. All firms draw a new potential price simultaneously. Without knowing each others' choices, they decide whether to stay or to switch, and they ignore that other firms are also changing their prices. As one would expect, volatility and dispersion are very high in this environment, because the entire set of firms updates simultaneously and myopically. It is perhaps less obvious that volatility and dispersion stay high even with a very large number of firms, in contrast to Model 1. In addition, a look at the autocorrelation function of profits (Figure 5) reveals a cyclical tendency in Model 2, and this tendency gets *especially* pronounced for large N.[8] Contrast this

[8] In Model 2 with $N = 100$, mean and variance of profits have a positive correlation of 0.8 (Table 2), and a negative autocorrelation peaking at a lag of 3 periods (Figure 5). There is no other way

Figure 4: **Chaos (profits and prices in Model 2)**



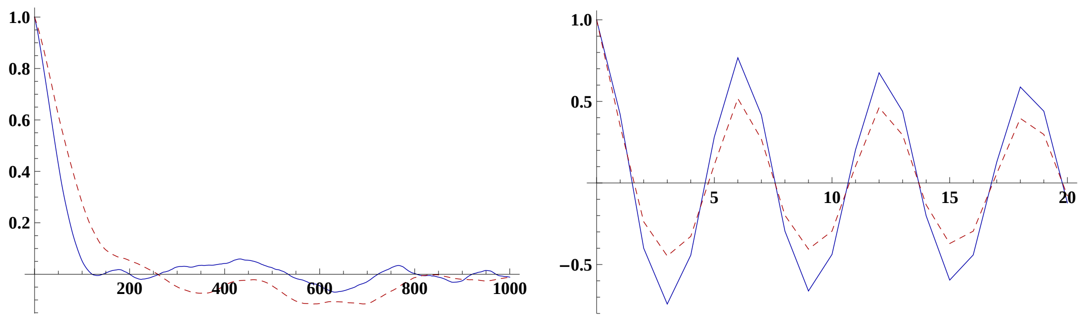
Cross-section mean and standard deviation (dashed) of profits, truncated to the first 1000 simulation periods.

Smoothed histograms of the cross-section mean (right) and standard deviation (dashed, left) of profits. Vertical lines indicate profit benchmarks π_∞ and π_N.

Density functions of prices pooled over all simulation periods, compared with the benchmark equilibrium (shaded area).

with Model 1 where autocorrelation is positive and declines to zero smoothly, as we would expect from updating in random sequential order.

Figure 5: **Autocorrelation in Models 1 and 2**



(a) Model 1 with $N = 100$ (b) Model 2 with $N = 100$

Autocorrelation functions of the mean (blue) and variance (red, dashed) of profits over simulated time. The difference in scale of the X-axes is mechanical: with 100 firms and sequential updating (Model 1) each firm gets a new draw every 100 periods on average. By contrast, in Model 2 all firms get a draw every period.

To be sure, the head-over-heels updating of Model 2 is even more artificial than the sequential updating of the other Models.[9] However, every firm updating at the same time is just an extreme representation of the idea that firms may not be able to observe their competitors' most recent actions when considering their own choices. And it is worth noting that profits are higher on average (albeit more volatile) in Model 2 than in Model 1, which suggests that even rational firm owners might not mind this myopic management strategy.

Results from Model 3

What are the distinguishing features of Burdett-Judd competition that generate these fragile results? There are two. First, if all firms have equal costs, the equilibrium requires indifference, and indifference is hard to learn for myopic decision makers with limited memory. Second, the best response function is not continuous: if all firms shifted their price up by ε, then the best response of a firm near the top of the

to describe this than that the profit distribution is moving like an accordion, played by a musician who is swaying to the rhythm of a waltz.

[9] None of the models satisfy rational expectations, as firms do not attempt to forecast future prices when switching, and they never learn that profit opportunities are ephemeral. But the discrepancy in Model 2 is particularly stark.

price ranking would not be to lower their price by a smooth function of ε, but it would be to move all the way to the bottom.[10]

But are both of these two features equally essential for fragility? We can look at this question by simulating a Burdett-Judd economy with cost dispersion. As Burdett and Judd (1983) proved for a continuum of firms with cost dispersion, firms with higher costs always charge higher prices, so the price ranking is identical to the cost ranking. In fact, if there is a non-atomic cost distribution, then a firm's individual choice problem becomes concave and determinate in equilibrium, even when the cost dispersion is tiny. But in the simulation world where decision making is inherently noisy, tiny may not be enough.

Table 3: **The correlation of costs and prices in Model 3**

$\bar{c}$	Average correlation	Standard deviation
1.1	0.47	0.21
1.4	0.83	0.08
1.9	0.94	0.03
Any continuum of firms with cost dispersion	1.00	0.00

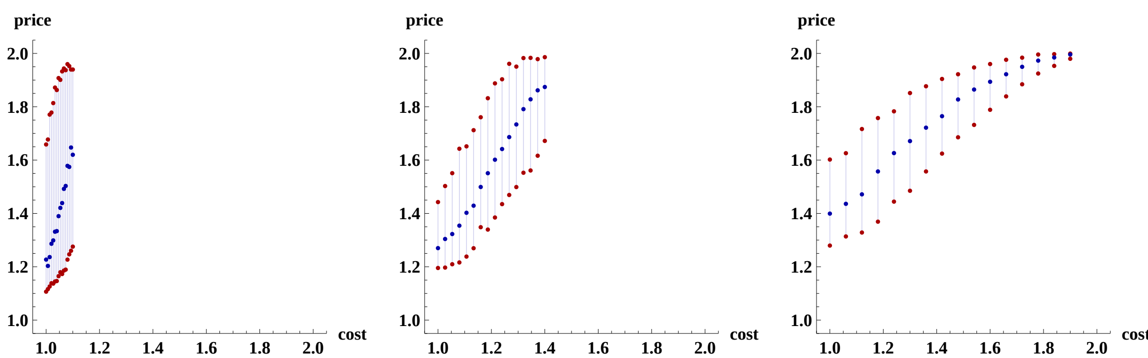
For each simulated time period, I compute Spearman's rank correlation coefficient for the cross-section of prices and costs, then report the time-series mean and standard deviation. There are $N = 16$ firms whose costs are evenly spaced between $c = 1$ and $\bar{c}$, and each firm faces the same unit elastic demand curve with reservation price $\bar{p} = 2$.

Model 3 implements such a variation of Model 1 with heterogeneous production costs. For concreteness, let $N = 16$. Each of the 16 firms is assigned a permanent cost, distributed uniformly over an interval. I simulate three cases: a narrow interval, an intermediate interval, and a broad interval where the highest cost in the market is almost equal to the reservation price. The results are summarized in Table 3 and illustrated in Figure 6.

[10] This distinguishes an equilibrium with endogenous price dispersion from simpler models of competition. For example, consider Cournot, where each firm's best response function varies smoothly with the actions of all other firms. I have also simulated this game with the same strategic frictions (the code and results are available upon request). In successive rounds of updating, firms do get closer and closer to the Cournot equilibrium. The noise dies down quickly for any N, and the rate of this dying-down is not related to N. Model 2, where myopia is taken to the extreme, actually converges faster than Model 1.

When cost dispersion is large, high-cost firms are forced to choose prices close to the reservation price, and thus low-cost firms charge low prices and take advantage of the associated higher sales. As a result, the price distribution tracks the cost distribution closely despite the noise introduced by sequential updating. As cost dispersion becomes small, on the other hand, the results resemble those of Model 1: all firms charge high prices at some times and low prices at others, and price-cost rank reversals are common.

Figure 6: **The distribution of costs and prices in Model 3**



(a) Costs between 1.0 and 1.1 (b) Costs between 1.0 and 1.4 (c) Costs between 1.0 and 1.9

Simulated percentiles of prices: 5th, 50th (median), and 95th. The intervals are not independent: when some firms charge relatively high prices, others are more likely to choose relatively low prices, and vice versa, which contributes to making price-cost rank reversals common.

We learn from this that the essential reason why the Burdett-Judd model is fragile is the shape of the best-response function, and it operates even when there is a theoretical benchmark that is fully determinate. In the basic Burdett-Judd model, the equilibrium profit function is *completely* flat; with a small amount of cost dispersion, it is still *quite* flat, and its curvature is not enough to overcome the fragility induced by the fact that every firm's action depends on everyone else's actions in a discontinuous way. We see this at work in the simulation: only when cost dispersion gets large enough does it reduce churn and improve the allocation of sales. ("Improve" in the following sense: with cost dispersion and constant returns to scale in producing a homogeneous good, it would be efficient to allocate all sales to the cheapest firm if there was a way to get around the search friction.)

Results from Model 4

In an environment where firms update their prices in order to grope their way towards strategic improvements – most of them small – it is natural to ask how the results are affected by the introduction of menu costs. In particular, in a "near-perfect" equilibrium where profit dispersion is small, a menu cost might be enough to lock in an equilibrium in which no firm wants to switch given the opportunity. And Figures 3 and 4 demonstrate that the dispersion of profits does indeed become very small on occasion. Consequently, I simulate selected cases of Models 1-3 with menu costs imposed.[11] The menu cost is only incurred if a firm chooses to change its price: the random draw of a new potential price is still observed for free. Table 4 summarizes the results.

We see that as menu costs increase, average profits and average profit dispersion decline in Models 1 and 2. As we saw earlier, profit levels and dispersion are positively correlated because outlier prices and profits tend to be "too high" rather than "too low"; menu costs tend to lock in low-dispersion states, and these are also the states with lower profits. Furthermore, volatility of profits increases in Model 1 and decreases in Model 2, making their results similar as menu costs get large. This is also sensible, because the only difference between them is whether updates are staggered or simultaneous. Menu costs reduce the frequency of updating – in the highest-cost case, they reduce the frequency of updating to 3% of all opportunities – which means that updates in Model 2 are now *effectively* staggered, too.

Looking at Model 3, we also see that in a sense, its performance "improves" as menu costs increase. In the Burdett-Judd benchmark, firms with higher costs must always charge higher prices, so the rank correlation should be 1 at all times. Indeed, this correlation gets close to 1 as menu costs increase. As updates become costly, firms only use them to correct particularly large price misalignments, and price-cost reversals become less common.

In the statistics above, profits are listed as gross, not subtracting menu costs paid. But we can compute adjustments and they are not large. In the middle cases ($d =$

[11] For ease of comparison, the menu costs are expressed as proportions of benchmark profits π_∞ for Models 1 and 2, and a slightly lower number, $\bar{\pi}_\infty$, for Model 3. The reason: with cost dispersion in the continuum benchmark, the highest-cost firm always charges the reservation price $\bar{p}$ and makes profits $\bar{\pi}_\infty = (1 - \bar{c}/\bar{p})J'(1)$, which differs from π_∞ in that the common cost c is replaced with the highest cost $\bar{c}$. Simulated profits will vary; for Models 1 and 2, most profits are between 1 and 1.5 times π_∞, and for Model 3, most profits are between 1 and 2 times $\bar{\pi}_\infty$.

Table 4: **Effect of menu costs**

Model	Menu cost d	Aggregate profit statistics			Update frequency
		Level	Volatility	Dispersion	
1	0	0.201	0.020	0.167	0.42
	$0.1\pi_\infty$	0.195	0.021	0.160	0.23
	$0.2\pi_\infty$	0.186	0.020	0.147	0.11
	$0.4\pi_\infty$	0.167	0.025	0.154	0.02
2	0	0.214	0.035	0.229	0.44
	$0.1\pi_\infty$	0.209	0.036	0.229	0.30
	$0.2\pi_\infty$	0.199	0.036	0.218	0.19
	$0.4\pi_\infty$	0.167	0.029	0.172	0.03
		Correlation of costs and prices			Update frequency
3 (with $\bar{c} = 1.4$)	0	0.83			0.30
	$0.1\bar{\pi}_\infty$	0.85			0.17
	$0.2\bar{\pi}_\infty$	0.88			0.10
	$0.4\bar{\pi}_\infty$	0.92			0.03

"Level", "Volatility", and "Dispersion" of profits are defined in Table 2. "Correlation" is defined in Table 3. "Update frequency" refers to a firm's choice to switch to a new price when given the opportunity. All runs use $N = 16$. Profit benchmarks are $\pi_{16} = 0.174$, $\pi_\infty = 0.167$, and $\bar{\pi}_\infty = 0.1$.

$\{0.1, 0.2\} \cdot \pi_\infty$), menu costs paid total only about 2-3% of profits across all runs. In the cases with the largest menu costs ($d = 0.4\pi_\infty$), updates decrease sufficiently that total menu costs paid go down, to about 1% of profits across the models.

In summary: we get closer to the continuum benchmark as menu costs increase, and menu costs seem to act in a pro-competitive way across all cases analyzed – even when the menu cost is so large that it eliminates nine out of ten updates that otherwise would have taken place. This supports the interpretation that in a noisy, second-best world, additional frictions can sometimes improve outcomes.

5 Conclusion

Models of endogenous price dispersion feature equilibria where firms are indifferent over a wide range of prices, and the best response function is not continuous even in equilibrium. This makes them different from other models of price setting, such as monopolistic competition, where there is generally a unique price which maximizes profits for each firm, and it depends smoothly on every other variable. If the number of firms is finite, it is impossible to satisfy the indifference condition exactly. This raises the question of how robust an endogenous price distribution can be in an environment of strategic frictions, and if a finite number of competitors would "find" it when starting from an arbitrary initial state.

Even if the current state almost perfectly satisfies indifference, two mechanisms will cause instability. First, a firm changing its price to take advantage of arbitrage opportunities will cause another arbitrage opportunity near the price it "vacated" in the price distribution, and the new one might be bigger than the original one. Second, if firms rush to update simultaneously, they may attempt to take advantage of profit opportunities which no longer exist.

In the simulation, I show that both mechanisms may contribute to price instability. Edgeworth cycles spontaneously arise and die out, as simulation excerpts show (Figure 2). Qualitatively, the benchmark comparison performs well when prices and profits are averaged over the long run, especially when updates are staggered. Quantitatively, however, differences persist even when the number of competitors is very large, and the differences are substantial when the number of competitors is small. Certainly, 6 or 16 competitors seems like a more realistic description of most markets than 100, even of markets for homogeneous goods like gas stations or retailers selling the same branded products. We should expect the fragility simulated in this paper to be at work in such markets.

One way to interpret the results is to appreciate that the simulated firms are extremely restricted in terms of their strategies: all they can do is myopic updating. Excepting Appendix A.2, they are not even able to compute potential profits over all possible prices. So, having neither memory nor foresight, we cannot expect them to "learn" a complex equilibrium. What is different, then, as N gets large? The difference is that with staggered updating (all Models except Model 2), each firm's price choice stays active for a while, and because these choices are not completely random, every updating firms' choice is guided by the past actions of all other firms. The market is *capable of memory* even though no individual firm is – as if it was guiding them with an invisible hand.

A Appendix

A.1 Comparing simulated and exact solutions for $N = 2$

For large but finite N, sampling with and without replacement each has pros and cons. With replacement, profits of an individual firm are tractable but the equilibrium price distribution is not. Without replacement, the price distribution is tractable (equal to F_∞, in fact, at least conditional on a fixed distribution of quotes), but profits of an individual firm are not. But for extremely low and high N, we do not have to choose and we can solve everything with pencil and paper. With $N \to \infty$, both protocols converge to the Burdett-Judd benchmark which was solved in the main text; at the other end, with $N = 2$, the issue of sampling boils down to the simple question which one of two prices is lower. Suppose the parameters are as before ($\bar{p} = 2, c = 1, \rho = 1, \eta = 2$), and we use sampling with replacement. We have $q_1 = 1/3$, so the most expensive firm would expect to make a fraction $1/(3N)$ of total sales if N was large. However, thanks to sampling with replacement, it actually makes more sales: a fraction $J(1) - J(1/2) = 1/4$ of the total when $N = 2$.

Suppose the two prices are $p_1 < p_2$ (ranking WLOG). Then realized profits are:

$$\pi(p_2) = \left(1 - \frac{1}{p_2}\right) \cdot \left[J(1) - J\left(\frac{1}{2}\right)\right] = \frac{1}{4}\left(1 - \frac{1}{p_2}\right)$$
$$\pi(p_1) = \left(1 - \frac{1}{p_1}\right) \cdot \left[J\left(\frac{1}{2}\right) - J(0)\right] = \frac{3}{4}\left(1 - \frac{1}{p_1}\right)$$

Using the standard Burdett-Judd indifference condition, we can find the mixed-equilibrium price distribution:

$$F_2(p) = \frac{3}{2} - \frac{1}{4}\frac{p}{p-1}$$

Expected profits are $1/8$ per firm, or $\pi_2 = 0.25$ total. We can also compute the analogues of "volatility", "dispersion", and "level-dispersion correlation" of profits. First, "volatility", the standard deviation of *total* profits over time, is expected to be:

$$\mathbb{E}\{\pi(p_1) + \pi(p_2)\} = \frac{1}{8}\sqrt{\frac{1}{2}\left[3\log^2(3) + 4(\log(9) - 3)\right]} \approx 0.057,$$

where p_1 and p_2 are distributed independently with CDF $F_2(p)$. And two measures of "dispersion", the standard deviation and the coefficient of variation of profits within a period, are expected to be:

$$\sigma_2 \equiv \mathbb{E}\left\{\frac{1}{2}\,|\pi(p_1)-\pi(p_2)|\right\} = \frac{1}{16}[4-3\log(3)] \approx 0.044$$

$$\mathbb{E}\left\{\frac{|\pi(p_1)-\pi(p_2)|}{\pi(p_1)+\pi(p_2)}\right\} = \frac{2\sigma_2}{\pi_2} \approx 0.352$$

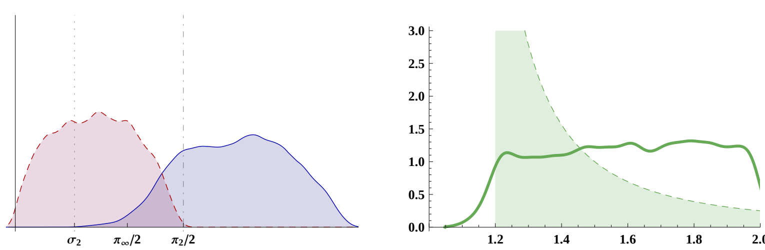
(The factor 2 appears because π_2 represents expected *total* profits, not the average per firm.) Either measure of dispersion turns out to be independent of the level of profits (hence the first equality in the last equation), so the "level-dispersion correlation" is expected to be zero.

Table A.1: **The case of** $N = 2$

	Level	Volatility	Dispersion	L-D correlation
Model 1 simulation	0.327	0.078	0.340	0.70
Mixed-strategy equilibrium	0.250	0.057	0.352	0.00

"Level" is the cross-sectional total of firms' profits in a simulated period, averaged across simulated time. "Volatility" is the time-series standard deviation of profit levels. "Dispersion" is the cross-sectional coefficient of variation of firms' profits in a simulated period, averaged across simulated time. The "L-D correlation" is the time-series correlation between the level of profits and their dispersion (measured by the coefficient of variation to avoid a mechanical correlation between means and standard deviations) in each simulated period.

Figure A.1: **The case of** $N = 2$



(a) Histograms of the cross-section mean (right, vs expectation $\pi_2/2$) and standard deviation (dashed, left, vs expectation σ_2) of profits.

(b) Histogram of prices, vs exact benchmark density $F_2'(p)$.

As Table A.1 and Figure A.1 show, the simulation of Model 1 with $N = 2$ yields something very far from the benchmark, though consistent with what one would expect from extrapolating Table 2 for small N. This confirms the main point of the paper: because it relies on a very complex strategic foundation, the Burdett-Judd framework may not be a good guide to competition in small markets, or in environments with strategic frictions. On the other hand, for large N and when strategic frictions are not too severe, the Burdett-Judd solution fits well.

A.2 Simulations with explicit profit maximization on a grid

The simulations in the main text had firms learn what their profits would be with one single, randomly picked price, and then decide whether to stay or to switch. One can imagine a different computational approach: firms choose prices from a discrete grid, maximizing profits over the whole grid when it is their turn to update (related to the "fictitious play" concept of Fudenberg and Kreps, 1993). I also simulated this alternative approach.

Model A.1

Firms choose prices freely from a discrete grid of 10 possible prices $\{p_i\}$, spaced equally between c (excluded) and $\bar{p}$ (included). In random sequential order, firms observe the existing price distribution; then, for each potential choice p_i, they compute the new price distribution if they did make that choice:

$$f_j(i) = \text{share of firms charging } p_j \quad \forall j = 1 \dots P$$
$$F_j(i) = \sum_{k=1}^{j} f_k \quad \forall j = 1 \dots P \quad \text{and} \quad F_0(i) = 0$$

and the implied profits, if they did make that choice:

$$\pi(i) = \left(1 - \frac{c}{p_i}\right) \frac{J(F_i(i)) - J(F_{i-1}(i))}{N f_i(i)}$$

Finally, they simply pick the price with the highest profits $\max_i\{\pi(i)\}$.

The goal of Model A.1 is to clarify whether the results of Model 1 (and, presumably, Models 2-4 as well) are an artifact of the way possible improvements are sampled. If Model A.1 gives similar results to Model 1, this suggests that the results are not an artifact of sampling, but rather are generic to strategic frictions in an environment sensitive to such frictions.

Model A.2

Same as Model A.1, but firms get their turn to update in fixed sequential order. Other than the initial condition, there is no randomness in this Model. The goal is to give deterministic dynamics a chance to emerge, in order to learn which ones do and which ones do not.

Models A.1 and A.2 are simulated for $N = 16$ and $N = 100$, with cost $c = 1$ for all firms. The other parameters are as in the main text: $\bar{p} = 2$, $\eta = 2$, and $\rho = 1$.

Table A.2: **Profits in Models A.1 and A.2**

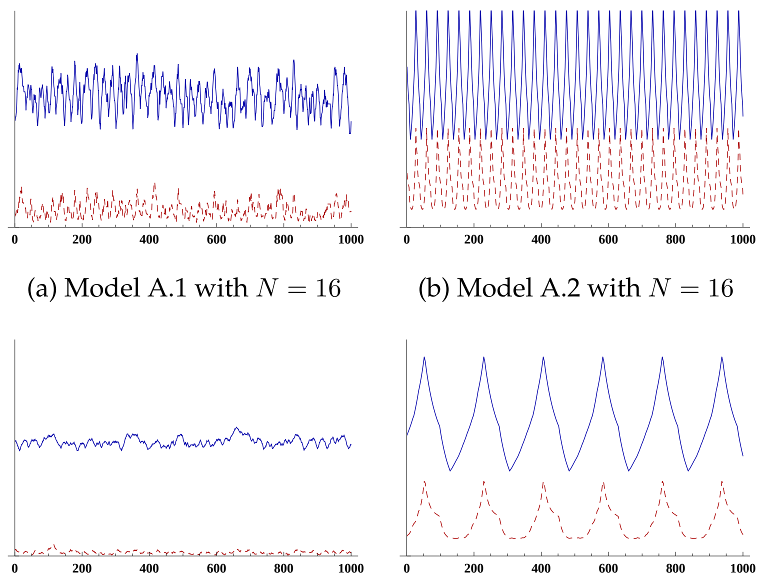
Model	N	Level	Volatility	Dispersion	L-D correlation
A.1	16	0.198	0.024	0.126	0.43
	40	0.184	0.013	0.068	0.35
	100	0.177	0.007	0.033	0.31
A.2	16	0.221	0.059	0.284	0.81
	40	0.207	0.053	0.258	0.81
	100	0.202	0.049	0.243	0.81
Continuum benchmark	∞	0.167	0.000	0.000	–

"Level" is the cross-sectional total of firms' profits in a simulated period, averaged across simulated time. "Volatility" is the time-series standard deviation of profit levels. "Dispersion" is the cross-sectional coefficient of variation of firms' profits in a simulated period, averaged across simulated time. The "L-D correlation" is the time-series correlation between the level of profits and their dispersion (measured by the coefficient of variation to avoid a mechanical correlation between means and standard deviations) in each simulated period.

The results depend crucially on the order in which firms update prices; see Table A.2 and Figure A.2. When the order is random, as in Model A.1, the results are almost identical to those obtained using Model 1 (see Table 2). In Model A.2 where the update order is fixed, however, every simulation run eventually reaches a stable cycle. For small N, the period of this pricing cycle is exactly two times the number of firms, implying that each firm ends up alternating between two prices. Consequently, the aggregate profit statistics of Model A.2 differ significantly from the 'well-behaved' models (1 and A.1), and they do not converge to the benchmark as the number of firms becomes large.

It is worth noting that among the many simulations run for this study, not a single one reached an absorbing state – in other words, a pure strategy Nash equilibrium of the simultaneous game – even though such equilibria exist when prices are restricted to a discrete grid.

Figure A.2: **Random vs fixed order of updating**



(a) Model A.1 with $N = 16$ (b) Model A.2 with $N = 16$

(a) Model A.1 with $N = 100$ (b) Model A.2 with $N = 100$

Time paths of the cross-section mean and standard deviation (dashed) of profits, truncated to the first 1000 simulation periods, for Models A.1 and A.2.

Does updating in fixed order necessarily lead to deterministic cycles? No, it does so only in Model A.2, where updating is fully deterministic. In Models 1-4, the updating procedure introduces additional randomness (firms can only see one new price, randomly selected, and can only decide whether to switch or stay). For these models, random versus fixed order updating does not make much of a difference.

References

Alessandria, G. (2009). Consumer search, price dispersion, and international relative price fluctuations. *International Economic Review 50*(3), 803–829.

Burdett, K. and K. L. Judd (1983, July). Equilibrium Price Dispersion. *Econometrica 51*(4), 955–69.

Burdett, K. and G. Menzio (2013, May). (q,s,s) pricing rules. Working Paper 19094, National Bureau of Economic Research.

Burdett, K. and D. T. Mortensen (1998, May). Wage Differentials, Employer Size, and Unemployment. *International Economic Review 39*(2), 257–73.

Burdett, K., A. Trejos, and R. Wright (2015). A simple model of monetary exchange with sticky and disperse prices. *mimeo*.

Cason, T. N., D. Friedman, and F. Wagener (2005). The dynamics of price dispersion, or edgeworth variations. *Journal of Economic Dynamics and Control 29*(4), 801–822.

Crockett, S., R. Oprea, and C. Plott (2011). Extreme walrasian dynamics: The gale example in the lab. *American Economic Review 101*(7), 3196–3220.

De los Santos, B., A. Hortaçsu, and M. R. Wildenbeest (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *The American Economic Review 102*(6), 2955–2980.

Fudenberg, D. and D. M. Kreps (1993). Learning mixed equilibria. *Games and Economic Behavior 5*(3), 320–367.

Geromichalos, A. (2014). Directed search and the bertrand paradox. *International Economic Review 55*(4), 1043–1065.

Head, A. and A. Kumar (2005, May). Price Dispersion, Inflation, And Welfare. *International Economic Review 46*(2), 533–572.

Head, A., A. Kumar, and B. Lapham (2010). Market power, price adjustment, and inflation. *International Economic Review 51*(1), 73–98.

Head, A., L. Q. Liu, G. Menzio, and R. Wright (2012). Sticky prices: A new monetarist approach. *Journal of the European Economic Association 10*(5), 939–973.

Herrenbrueck, L. (2015). An open-economy model with money, endogenous search, and heterogeneous firms. Working paper, Simon Fraser University.

Hopkins, E. and R. M. Seymour (2002). The stability of price dispersion under seller and consumer learning. *International Economic Review 43*(4), 1157–1190.

Kaplan, G. and G. Menzio (2016). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy 124*(3), 771–825.

Kaplan, G., G. Menzio, L. Rudanko, and N. Trachter (2016). Relative price dispersion: evidence and theory. Technical report, National Bureau of Economic Research.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1996). *Microeconomic theory*. Oxford University Press New York.

Maskin, E. and J. Tirole (1988, May). A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles. *Econometrica 56*(3), 571–99.

Menzio, G. and N. Trachter (2015). Equilibrium price dispersion across and within stores. Technical report, National Bureau of Economic Research.

Mortensen, D. T. (2005, May). A Comment On "Price Dispersion, Inflation, And Welfare" By A. Head And A. Kumar. *International Economic Review 46*(2), 573–578.

Wang, L. (2016). Endogenous search, price dispersion, and welfare. *Journal of Economic Dynamics and Control 73*, 94–117.