Private and Social Learning in Frictional Product Markets*

Sephorah Mangin
ANU

Guido Menzio
NYU and NBER

September 4, 2024

Abstract

We introduce dynamic learning in the static search-theoretic framework of imperfect competition of Butters (1977), Varian (1980) and Burdett and Judd (1983). We consider two forms of learning: memory and word of mouth. In the model with memory, long-lived buyers not only learn about some new sellers in every period, but they also remember some of the sellers about which they learned in the past. In the model with word of mouth, short-lived buyers not only learn about sellers through search, but they also learn about sellers by talking to previous buyers, which refer them to the best seller of which they are aware. Both models are tractable. We establish the existence and uniqueness of equilibrium and characterize the dynamics of the product market. Memory increases the quantity of information available to buyers and, for this reason, leads to higher concentration and more competition over time for all sellers. Word of mouth increases the quality of information available to buyers and leads to higher concentration over time, but it does not increase competition for all sellers. With word of mouth, the dynamics of competition are heterogeneous across sellers and depend on a seller's popularity. For different sellers, competition and prices may be increasing, decreasing, or non-monotonic over time.

JEL Codes: D43, D83, J31.

Keywords: Search frictions, Imperfect competition, Market dynamics, Memory, Word of mouth.

*Sephorah Mangin: Research School of Economics, Australian National University, HW Arndt Building 25A, Canberra, ACT 2601 (email: sjmangin@gmail.com). Guido Menzio: Department of Economics, New York University, 19 West 4th Street, New York, NY 10012 (email: gm1310@nyu.edu).

1 Introduction

This paper contributes to the development of the search-theoretic model of imperfect competition in product markets of Butters (1977), Varian (1980), and Burdett and Judd (1983). The simple idea behind the search-theoretic model of imperfect competition is that, due to information or physical frictions, buyers cannot purchase from any seller in the market but only from a discrete subset of sellers. Since buyers have a limited choice set, the equilibrium of the market is imperfectly competitive – where the extent of actual competition depends on the distribution of the size of buyers' choice sets. The model has been fruitfully applied to study price dispersion (e.g., Sorensen 2000, Galenianos et al. 2012, Kaplan and Menzio 2015, Kaplan et al. 2019, Shi 2023), price stickiness (Head et al. 2012, Burdett and Menzio 2018), monetary exchange (Head and Kumar 2005, Burdett et al. 2017), the variation of markups across groups of buyers (Pytka 2017, Nord 2023) and over time (Kaplan and Menzio 2016), specialization patterns (Menzio 2023), store locations (Cai et al. 2023), and market power (Menzio 2024). These applications, even those that are explicitly dynamic, assume that the quantity and the quality of the information possessed by buyers about sellers does not improve over time. The assumption keeps the model static and the analysis of equilibrium simple. The assumption, however, may not be a good description of product markets in which buyers are long-lived, or markets in which buyers can gather information from others in the market.

In this paper, we contribute to the development of the search-theoretic framework of imperfect competition by introducing *learning*. We consider two forms of learning: memory and word of mouth. In the model with memory, we consider a version of the framework in which long-lived buyers not only learn about some new sellers in every period, but they also remember some of the sellers that they contacted in the past. As a result of this process of private learning, the quantity of information that buyers have about sellers increases over time. In the model with word of mouth, we consider a version of the framework in which short-lived buyers not only learn about sellers through search, but they also learn about sellers from talking to previous buyers, which refer them to the best seller of which they are aware. As a result of this process of social learning, the quality of the information that buyers have about sellers increases over time. Both types of learning lead to non-trivial, albeit different, market dynamics.

In Section 2, we consider a version of the search-theoretic model of imperfect competition in which buyers have memory. Specifically, we consider a product market that is populated by a continuum of long-lived sellers and by a double continuum of long-lived buyers. Sellers are heterogeneous with respect to the quality of their variety of the product. In every period, sellers post prices. Buyers are ex-ante homogeneous and demand one unit of the product per period. In every period, buyers search the market and contact m new sellers, where m is distributed as a Poisson with some coefficient μ. Buyers also have a probability $1 - \delta$ of remembering a seller that they had contacted in the past.

We show that equilibrium exists and is unique. The dynamics of equilibrium are driven by the evolution of the average number of sellers λ_t in the choice set of buyers, which is equal to $(1-\delta)\lambda_{t-1}+\mu$, the average number of remembered sellers $(1-\delta)\lambda_{t-1}$ plus the average number of new sellers μ a buyer contacts. Since λ_t is strictly increasing over time, the choice set of buyers expands and the market becomes progressively more competitive. The increase in competition allows buyers to purchase the good from sellers with higher quality, since sellers with higher quality offer higher surplus to buyers. The increase in competition also induces sellers to lower prices. While competition increases over time as the average size of buyers' choice set expands, the market does not become perfectly competitive as long as buyer's memory is imperfect. Intuitively, as long as memory is imperfect, each seller has a strictly positive probability of meeting a captive buyer and, hence, can always secure itself strictly positive profits. We also show that prices are lower with better memory, i.e. the price distribution in any given period is strictly decreasing, in the sense of first-order stochastic dominance, with respect to the buyers' memory $1-\delta$.

In Section 3, we consider a version of the search-theoretic model of imperfect competition with word of mouth. Specifically, we consider a product market that is populated by a continuum of long-lived sellers and by a double continuum of short-lived buyers. Sellers are heterogeneous with respect to the quality of their variety of the product. In every period, sellers post prices. Buyers are homogeneous and demand one unit of the good. In every period, buyers search the market and contact m sellers, where m is distributed as a Poisson with coefficient μ. Buyers also meet r old buyers, where r is distributed as a Poisson with coefficient ρ. When new and old buyers meet, the old buyers tell the new buyers about the best seller of which they were aware in the previous period.

We show that equilibrium exists and is unique. By contrast to the model with memory, the average number of sellers in the choice set of buyers is constant over time and equal to $\rho+\mu$, the average number of word-of-mouth contacts ρ plus the average number of new sellers μ a buyer contacts. The dynamics of equilibrium are driven entirely by the evolution of the *sampling distribution*, i.e. the distribution of sellers from which buyers sample. Buyers active in period t sample some sellers through search from the distribution of all sellers, but they sample additional sellers through word of mouth from a more selective distribution: the distribution of the highest quality seller known to buyers who purchased in period $t-1$. Buyers active in period t purchase the good from the highest quality seller among those contacted through either search or word of mouth.

Along the equilibrium path, the distribution of the highest quality seller known to buyers who purchased in the previous period gets better and better over time. Therefore, over time, the size of the buyers' choice sets does not increase, but its quality does. The increase in the quality of the buyers' choice sets leads to higher sales concentration over time —just as memory does. However, the dynamics of competition are more complex with word of mouth. With word of mouth, competition does not increase over time for all sellers, as it does with memory. The dynamics of competition are heterogeneous across

sellers and depend on the seller's *popularity* – which is a measure of how likely it is that the seller will be contacted through word of mouth rather than search. We find that competition may be either increasing or decreasing, or it may vary non-monotonically over time, depending on a seller's popularity. In any given period of time, the highest quality sellers face increasing competition, but the lowest quality sellers face decreasing competition. Overall, the increase in the quality of buyers' choice sets does not lead to lower prices to the same degree as memory. Specifically, we show that, for the same level of sales concentration, prices in the model with word of mouth are always higher, for any given quality y, than prices in the model with memory.

In the Online Appendix, we present a version of the memory model in which buyers enter and exit the market. The properties of equilibrium described in Section 2 carry over when there is buyer entry and exit. In addition, this version of the model allows us to study the dynamics of individual buyers independently from the dynamics of the market. There is a distribution of buyers' *experience*, i.e. the number of periods a buyer has been in the market, which evolves over time. We show that the surplus captured by a buyer is higher, in expectation, with more experience in the market. We also show that a buyer purchases from higher quality sellers, in expectation, with more experience.

Related literature. As discussed, this paper introduces dynamic learning into the search-theoretic model of imperfect competition of Butters (1977), Varian (1980), and Burdett and Judd (1983). Our introduction of both private learning through memory and social learning through word-of-mouth communication contributes to the large search literature that applies and extends this model including Head and Kumar (2005), Head et al.(2012), Kaplan and Menzio (2015), Kaplan and Menzio (2016), Pytka (2017), Burdett and Menzio (2018), Kaplan et al. (2019), Nord (2023), Menzio (2023), Cai et al. (2023), and Menzio (2024). Our model with learning through word of mouth is also related to the search-theoretic literature on referral networks in the labor market, such as Galenianos (2013, 2014, 2021). Our model provides a highly tractable dynamic framework for capturing both search frictions and learning under imperfect competition.

This paper also contributes to the literature on word-of-mouth communication in product markets, such as Galeotti (2010), Campbell (2013, 2015, 2019), Campbell et al. (2020), and Campbell et al. (2024).[1] This literature often considers either a monopoly (Campbell 2013, 2015) or a (potential) duopoly (Campbell et al. 2020 and Campbell et al. 2024) and studies how *awareness* of a firm's product changes over time due to word of mouth and how this affects demand, entry, market shares or market concentration. Our approach is complementary but different. By using the search-theoretic model of imperfect competition, we obtain results regarding the effects of word of mouth on the dynamics of competition in an environment where prices are determined by local oligopoly. In particular, we show how these dynamics differ across low and high-quality sellers.

[1]Early papers on word-of-mouth learning in general environments include Ellison and Fudenberg (1995) and Banerjee and Fudenberg (2004).

2 Memory

In this section, we construct a version of the search-theoretic model of Butters (1977), Varian (1980) and Burdett and Judd (1983) in which buyers can remember sellers about which they learned in the past. We construct the model so that, even though buyers and sellers are long-lived, the buyer's problem of choosing where to purchase the good is static, and the seller's problem of choosing what price to post is static. We establish that the equilibrium exists and is unique and we characterize the evolution of equilibrium outcomes over time.

2.1 Environment

We consider the market for some consumer good. On one side of the market, there is a measure 1 of infinitely-lived sellers. Sellers are heterogeneous with respect to the quality y of their variety of the good.[2] The distribution of sellers with respect to y is given by a twice differentiable cumulative distribution function $\Phi(y)$ with $\Phi'(y) > 0$ and support $[y_\ell, y_h]$ where $0 < y_\ell < y_h$. A seller produces its variety of the good at a constant marginal cost, which we set to 0 for the sake of simplicity.[3] In every period $t = 1, 2, ...$, the seller posts a price p for its variety of the good.[4] If the seller trades q units of the good at the price p, its periodical profit is qp.

On the other side of the market, there is a continuum of measure b of infinitely-lived buyers per seller.[5] Buyers are ex-ante homogeneous. In every period t, a buyer demands

[2]We assume that sellers are heterogeneous with respect to the quality of their variety of the product. In a version of the model with homogeneous sellers, the price posted by an individual seller is indeterminate, since the equilibrium involves sellers mixing over prices according to an equilibrium distribution. In a version of the model with heterogeneous sellers, the price posted by an individual seller is uniquely pinned down. Heterogeneity purifies mixed strategies. The reader is free to interpret the heterogeneity in the model as a purification device, in which case the reader can take the extent of heterogeneity to be arbitrarily small, or as economically meaningful heterogeneity, in the sense that some sellers carry a better variety of the product than others. Supporting the second interpretation, Albrecht, Menzio and Vroman (2023) show that, if sellers can invest in the quality of their product, the unique equilibrium of the model features heterogeneity.

[3]The assumption that the marginal cost of production is 0 is made for the sake of algebraic simplicity, and all the results presented in the paper extend trivially to the case in which the marginal cost of production is strictly positive. The assumption that the marginal cost of production is constant, rather than strictly increasing, is substantive. Menzio (2023) shows that the structure of equilibrium is qualitatively different when sellers operate a production function with decreasing rather than constant returns to scale.

[4]We assume that a seller can only post a price. A seller cannot post a schedule of prices in an attempt to discriminate buyers. Specifically, a seller is not allowed to post a price schedule $p(h_t)$, where the price paid by the buyer depends on its purchasing history h_t. The assumption seems natural in the context of retail markets.

[5]We assume that the market is populated by a continuum of sellers and a double continuum of buyers. If we were to assume that the market is populated by a continuum of sellers and a continuum of buyers, the actual number of buyers that meet an individual seller and, in turn, the actual number of trades made by an individual seller would be a random variable. The realization of the random variable would be informative about the choice set of the buyers that purchased the good from the seller and, in turn, the seller would use this information to set its price in the next period. For instance, if the seller made more trades than expected, the seller would infer that it is in contact with a larger than expected number

one unit of the good. If the buyer purchases a variety of the good of quality y at the price p, its periodical utility is $y - p$. If the buyer does not purchase the good, its periodical utility is 0.

The product market is frictional, in the sense that a buyer cannot purchase the good from any seller in the market, but only from those sellers with which it is in contact. In period 1, buyers enter the market without any contacts. At the beginning of each period $t > 1$, a buyer is in contact with n_{t-1} sellers. For any given one of the n_{t-1} sellers, the buyer permanently loses contact with that seller with probability $\delta \in [0, 1]$. During period t, the buyer contacts m_t additional sellers, where $m_t = 0, 1, 2, \ldots$ is drawn from a Poisson distribution with coefficient $\mu > 0$. At the end of period t, the buyer chooses whether and where to purchase the good among the sellers with which it is in contact.

The environment described above is quite natural. In every period, buyers search the market and discover some new sellers. Buyers are also aware of the sellers that they have discovered in the past. The ability of buyers to remember sellers that they have discovered in the past is limited by the fact that a buyer "forgets" about a previously discovered seller with probability δ. If $\delta = 1$, buyers have no memory, and the environment is the same as in Butters (1977), Varian (1980) and Burdett and Judd (1983). If $\delta = 0$, buyers have perfect memory. If $\delta \in (0, 1)$, buyers have some memory, but their memory is imperfect.

The parameter δ may be interpreted literally as the probability that a buyer forgets about a seller. The parameter δ may also capture the probability of a negative shock to a time-varying, match-specific component of the gains from trade between the buyer and the seller that drives the gains from trade to zero. Notice that buyers have the same probability of "forgetting" a seller whether or not they purchased the good from it. This assumption is the polar opposite of the assumption made in the search model of imperfect competition in the labor market of Burdett and Mortensen (1998). Indeed, in this model, workers are assumed to have no memory of any of the previously contacted firms, except for the firm at which they are currently employed. This assumption greatly simplifies the analysis of equilibrium as it makes the buyer's choice of where to purchase the good static.

2.2 Equilibrium

In order to characterize equilibrium, we start by analyzing the properties of the buyer's choice set. At the beginning of period 1, a buyer enters the market without any contacts. During period 1, the buyer searches the market and comes into contact with m_1 sellers, where m_1 is a Poisson random variable with coefficient μ. Hence, at the end of period

of buyers for which the seller is the most attractive choice. Since the identity of the sellers in the buyers' choice sets is persistent, the seller could use this information to set its price in the next period. In contrast, under the assumption of a double continuum of buyers, the measure of buyers that meet an individual seller is deterministic and so are the measure of trades made by the seller. Therefore, the seller cannot infer anything about buyers from observing the quantity of the good that it has traded.

1, the buyer can purchase the good from $n_1 = m_1$ sellers. At the beginning of period 2, the buyer is in contact with n_1 sellers. During period 2 and for each one of the n_1 sellers, the buyer has a probability δ of losing contact with the seller, and a probability $1 - \delta$ of staying in contact with the seller. Let $\hat{n}_2$ denote the number of sellers with which the buyer remains in contact. During period 2, the buyer searches the market and comes into contact with m_2 additional sellers, where m_2 is a Poisson random variable with coefficient μ. Therefore, at the end of period 2, the buyer can purchase the good from $n_2 = \hat{n}_2 + m_2$ sellers. The buyer keeps adding and losing contacts in the same way in period $t = 3, 4,...$

By assumption, the number of buyer's contacts n_1 is distributed as a Poisson with coefficient μ. The next lemma establishes that the number of a buyer's contacts n_t is distributed as a Poisson also in any period $t > 1$. There is a simple logic behind this result. If the buyer enters period t with a number of contacts n_{t-1} that is Poisson with coefficient λ_{t-1} and each contact is maintained with probability $1 - \delta$, the number of contacts $\hat{n}_t$ maintained by the buyer is distributed as a Poisson with coefficient $(1 - \delta)\lambda_{t-1}$. In turn, if the number of contacts $\hat{n}_t$ maintained by the buyer is distributed as a Poisson with coefficient $(1 - \delta)\lambda_{t-1}$ and the buyer contacts a number m_t of additional sellers that is distributed as a Poisson with coefficient μ, the buyer's number n_t of contacts at the end of period t is Poisson with coefficient λ_t, where λ_t is given by $\mu + (1 - \delta)\lambda_{t-1}$.

It is clear from Lemma 1 that λ_t is strictly increasing over time if $\delta \in (0, 1)$.

Lemma 1: *The number n_t of buyer's contacts in period $t = 1, 2, \ldots$ is distributed as a Poisson with coefficient λ_t, where* λ_t is given by

$$\lambda_t = \frac{\mu}{\delta}[1 - (1 - \delta)^t]. \tag{2.1}$$

Having characterized the properties of the buyer's choice set, we can now examine the problem of a buyer. In period t, the buyer is in contact with n_t sellers. For each of these sellers, the buyer observes the quality y of their variety of the product and the price p that they charge. The buyer's decision affects its payoff in the current period. In particular, if the buyer purchases a variety of the good of quality y at the price p, the buyer enjoys a payoff of $y - p$ in the current period. If the buyer does not purchase the good, the buyer enjoys a payoff of 0 in the current period. The buyer's decision does not affect the buyers' future payoffs. The buyer's decision does not affect the buyer's contacts in the future (since the buyer is equally likely to remember all sellers they contact), nor does it affect the buyer's prices in the future (since sellers are not allowed to discriminate based on purchasing history). The buyer's problem is static, and it simply involves comparing the surplus $s = y - p$ offered by the n_t sellers and the payoff from not purchasing the good. Obviously, the solution to the buyer's problem is such that the buyer purchases from the seller that offers the highest surplus s if such surplus is positive, and the buyer does not purchase the good if s is strictly negative. If the highest positive surplus s is offered by multiple sellers, the buyer purchases the good from any of them with equal probability.

We now turn to examine the problem of a seller with a variety of quality y in period t. The seller has to decide the price p for its variety of the good or, equivalently, the surplus $s = y - p$ offered to buyers. The seller's decision has no effect on the seller's demand in the future, since the measure and the distribution of buyers that come into contact with the seller in the future is independent of the seller's current price. The seller's decision has no effect on the seller's ability to set prices in the future, since the current price does not restrict the seller's future prices. Therefore, the seller's problem is static, and it amounts to choosing s to maximize profits in period t.

For any $s \geq 0$, the seller's profit in period t is given by

$$V_t(y, s) = \left[\sum_{k=0}^{\infty} b_{k,t}\pi_{k,t}(s)\right](y - s), \tag{2.2}$$

where

$$b_{k,t} = b\frac{e^{-\lambda_t}\lambda_t^{k+1}}{(k+1)!}(k+1), \tag{2.3}$$

and

$$\pi_{k,t}(s) = F_t(s-)^k + \sum_{j=1}^{k} \frac{j!(k-j)!}{k!}\frac{\zeta_t(s)^j F_t(s-)^{k-j}}{(j+1)}. \tag{2.4}$$

Let us explain the expressions above. The seller meets a measure $b_{k,t}$ of buyers that are in contact with k other sellers, with $k = 0, 1, 2 \ldots$ The measure of buyers $b_{k,t}$ who have k other contacts is given by (2.3), which is the measure of buyers that have $k+1$ contacts, given that the number of contacts of each buyer is Poisson with coefficient λ_t, multiplied by $k+1$, the number of contacts for each one of these buyers. A buyer with k other contacts purchases from the seller with probability $\pi_{k,t}(s)$. The probability $\pi_{k,t}(s)$ is given by (2.4), where $F_t(s)$ denotes the cumulative distribution function of surplus offered by sellers, $F_t(s-)$ denotes the left limit of F_t and is the fraction of sellers offering strictly less than s, and $\zeta_t(s)$ denotes the fraction of sellers that offer s. The probability $\pi_{k,t}(s)$ is the sum of the probability of two events. The first event is that the buyer's k other contacts offer a surplus strictly smaller than s. The second event is that j of the buyer's k other contacts offer a surplus equal to s and $k - j$ of them offer a surplus strictly smaller than s and the buyer breaks the indifference in favor of the seller. The seller's periodical profit in (2.2) is given by the number of buyers that purchase from the seller multiplied by the seller's profit per unit sold.

The next three lemmas are standard fare in the analysis of models in the style of Butters (1977), Varian (1980), and Burdett and Judd (1983). The first lemma establishes that the profit of the seller is strictly positive. The second lemma establishes that the surplus distribution $F_t(s)$ cannot have any mass points. The third lemma establishes that the support of the surplus distribution $F_t(s)$ is an interval $[s_{\ell,t}, s_{h,t}]$, with $s_{\ell,t} = 0$.

Lemma 2: *For any $y \in [y_\ell, y_h]$, the seller's maximized profit is strictly positive.*

Lemma 3: *The equilibrium surplus distribution $F_t(s)$ does not have mass points.*

Since the surplus distribution does not have any mass points, $\zeta_t(s) = 0$ and $F_t(s-) = F_t(s)$ for all s. Using these observations, we can simplify (2.2) as

$$\begin{aligned} V_t(y,s) &= \left[\sum_{k=0}^{\infty} b\frac{e^{-\lambda_t}\lambda_t^{k+1}}{(k+1)!}(k+1)F_t(s)^k\right](y-s) \\ &= b\lambda_t e^{-\lambda_t}e^{\lambda_t F_t(s)}\left[\sum_{k=0}^{\infty}\frac{e^{-\lambda_t F_t(s)}\lambda_t^k F_t(s)^k}{k!}\right](y-s) \\ &= b\lambda_t e^{-\lambda_t(1-F_t(s))}(y-s). \end{aligned} \tag{2.5}$$

The first line is obtained by substituting $b_{k,t}$ and $\pi_{k,t}(s)$ from (2.3) and (2.4) into (2.2), and by using the fact that $\zeta_t(s) = 0$ and $F_t(s-) = F_t(s)$. The second line is obtained by collecting terms. The third line is obtained by recognizing that the summation in the second line is equal to 1. In expression (2.5) for $V_t(y,s)$, the term $b\lambda_t e^{-\lambda_t(1-F_t(s))}$ is the measure of buyers such that the surplus s offered by the seller is the highest surplus offered to those buyers, i.e. it is the seller's trade volume.

Lemma 4: *The support of the equilibrium surplus distribution F_t is an interval $[s_{\ell,t}, s_{h,t}]$, with $s_{\ell,t} = 0$.*

The next lemma establishes that the surplus offered by a seller is a strictly increasing function of the seller's quality. The finding is intuitive. The benefit of increasing the surplus s offered to buyers—which is to the increase the quantity of the good sold—is strictly greater for a seller with a variety of the good of higher rather than lower quality, since the profit per sale enjoyed by a seller is strictly increasing in y.

Lemma 5: *In equilibrium, the surplus offered by a seller is a strictly increasing function $s_t(y)$ of its quality.*

Given that buyers choose the highest surplus among the sellers they contact, the fact that $s_t(y)$ is strictly increasing in quality y means that the distribution of qualities across transactions is equal to the distribution of the highest quality seller a buyer contacts, conditional on the buyer having at least one contact. Letting $H_t(y)$ denote the fraction of transactions at sellers with quality smaller than y, we have

$$\begin{aligned} H_t(y) &= \left[\sum_{k=1}^{\infty}\frac{e^{-\lambda_t}\lambda_t^k}{k!}\Phi(y)^k\right]\Big/\left[1-e^{-\lambda_t}\right] \\ &= \left[e^{-\lambda_t}\sum_{k=0}^{\infty}\frac{(\lambda_t\Phi(y))^k}{k!} - e^{-\lambda_t}\right]\Big/\left[1-e^{-\lambda_t}\right] \\ &= \frac{e^{-\lambda_t(1-\Phi(y))} - e^{-\lambda_t}}{1-e^{-\lambda_t}}. \end{aligned} \tag{2.6}$$

The numerator in the first line of (2.6) is the product of the probability that a buyer has k contacts and the probability, $\Phi(y)^k$, that all of its k contacts have quality less than y,

summed over $k = 1, 2, \ldots$. The denominator in the first line is the probability that the buyer has at least one contact and, hence, purchases the good. The third line of (2.6) uses the fact that $\sum_{k=0}^{\infty} x^k/k! = e^x$.

The transaction distribution $H_t(y)$ first-order stochastically dominates the distribution of seller quality, i.e. $H_t(y) < \Phi(y)$. The average quality of a transaction, $\tilde{y}_t = E_{H_t}(y)$, is given by

$$\tilde{y}_t = \frac{\int x \lambda_t \Phi'(x) e^{-\lambda_t(1-\Phi(x))} dx}{1 - e^{-\lambda_t}}. \tag{2.7}$$

We can similarly define the average surplus per transaction, $\tilde{s}_t = E_{H_t}(s_t(y))$, and the average price paid per transaction, $\tilde{p}_t = E_{H_t}(p_t(y))$.

Returning to the seller's problem, a seller producing quality y solves:

$$V_t(y) = \max_{s \in [0,y]} b\lambda_t e^{-\lambda_t(1-F_t(s))}(y - s). \tag{2.8}$$

The necessary condition for the optimality of the surplus $s_t(y)$ offered by a seller of quality y is

$$b\lambda_t^2 e^{-\lambda_t(1-F_t(s_t(y)))} F_t'(s_t(y))(y - s_t(y)) = b\lambda_t e^{-\lambda_t(1-F_t(s_t(y)))}. \tag{2.9}$$

The left-hand side of (2.9) is the marginal benefit of offering buyers an additional unit of surplus, which is given by the increase in the volume of sales multiplied by the profit per sale. The right-hand side of (2.9) is the marginal cost of offering buyers an additional unit of surplus, which is given by the volume of sales. The optimality condition (2.9) states that the marginal benefit of offering an additional unit of surplus must be equal to the marginal cost.

The surplus distribution F_t is such that

$$F_t(s_t(y)) = \Phi(y). \tag{2.10}$$

The left-hand side of (2.10) is the fraction of sellers that offer surplus smaller than $s_t(y)$. The right-hand side of (2.10) is the fraction of sellers with a quality smaller than y. Since $s_t(y)$ is strictly increasing function, the left and the right-hand sides of (2.10) must be equal for every y. Differentiating (2.10) with respect to y implies that $F_t'(s_t(y))s_t'(y)$ is equal to $\Phi'(y)$. Combining this observation with (2.10) allows us to rewrite (2.9) as

$$s_t'(y) = \Phi'(y)\lambda_t(y - s_t(y)). \tag{2.11}$$

The expression above is a differential equation for the surplus function $s_t(y)$. The relevant solution of the differential equation satisfies the boundary condition $s_t(y_\ell) = 0$, since Lemma 4 guarantees that the lowest surplus on the distribution F_t is $s_{\ell,t} = 0$, and Lemma 5 guarantees that the seller that offers the lowest (highest) surplus on the distribution F_t is the seller with the lowest (highest) quality, i.e. $s_t(y_h) = s_{h,t}$ and $s_t(y_\ell) = s_{\ell,t}$.

The analysis above identifies a unique candidate equilibrium, which is given by the

solution to the differential equation (2.11) together with the boundary condition $s_t(y_\ell) = 0$. To verify that the candidate equilibrium is an equilibrium, consider a seller with an arbitrary quality $y_0 \in [y_\ell, y_h]$. In the candidate equilibrium, the seller offers the surplus $s_t(y_0)$, where $s_t(y_0)$ satisfies the optimality condition (2.9). In the candidate equilibrium, the seller makes a strictly positive profit. The seller does not want to deviate and offer a surplus $\hat{s}$ that is strictly negative, since by doing so its profit would be zero. The seller does not want to deviate and offer a surplus $\hat{s} \in [0, s_t(y_0))$. In fact, $\hat{s} = s_t(\hat{y})$ for some $\hat{y} < y_0$ and, hence, the derivative of the seller's objective function (2.9) is equal to zero for $\hat{y}$, and strictly positive for y_0. Similarly, the seller does not want to deviate and offer a surplus $\hat{s} \in (s_t(y_0), s_t(y_h)]$. Finally, the seller does not want to deviate and offer a surplus $\hat{s} > s_t(y_h)$, since in doing so it would attain a profit that is strictly smaller than by offering the surplus $s_t(y_h)$.

We have thus established the existence and uniqueness of equilibrium.

Theorem 1: (Existence and uniqueness of equilibrium with memory) *An equilibrium exists and is unique. For $t = 1, 2, ...$, equilibrium is given by a surplus function $s_t(y)$ that satisfies the differential equation (2.11) together with the boundary condition $s_t(y_\ell) = 0$. The solution to this differential equation is*

$$s_t(y) = \frac{\int_{y_\ell}^{y} \lambda_t e^{-\lambda_t(1-\Phi(x))} x \Phi'(x) dx}{e^{-\lambda_t(1-\Phi(y))}}, \tag{2.12}$$

and $s_t(y_h) = (1 - e^{-\lambda_t})\tilde{y}_t$ where $\tilde{y}_t$ is the average transaction quality.

To interpret this expression, let $\hat{H}_t(y) = e^{-\lambda_t(1-\Phi(y))}$, the distribution of the highest quality among the sellers a buyer meets.[6] Expression (2.12) is equivalent to

$$s_t(y) = \frac{\int_{y_\ell}^{y} x \hat{H}_t'(x) dx}{\hat{H}_t(y)}.$$

The surplus $s_t(y)$ is equal to the expected value of the highest quality among the sellers a buyer contacts, conditional on that quality being less than the seller's own quality y. This is equivalent to the expected value of the second-best seller's quality among the sellers a buyer contacts, conditional on y being the highest seller's quality. Effectively, firms set prices $p_t(y) = y - s_t(y)$ equal to the expected value of the difference between their own quality y and the next best quality among the sellers a buyer meets.

Associated with the equilibrium surplus function $s_t(y)$, there are a surplus distribution $F_t(s)$ and a price distribution $G_t(p)$. The surplus distribution $F_t(s)$ is given by

$$F_t(s) = \Phi(y_t(s)), \tag{2.13}$$

where $y_t(s)$ is a strictly increasing function that denotes the inverse of the surplus function

[6]Note that this distribution has a mass point at y_l because we assume that if the buyer meets no sellers the maximum is y_l, and thus the buyer gets surplus $s(y_l) = 0$.

$s_t(y)$. Similarly, the distribution of prices posted by sellers, $G_t(p)$, is given by

$$G_t(p) = \int_{y:p_t(y)\leq p} \Phi'(y)dy, \tag{2.14}$$

where $p_t(y)$ is the price function defined as $p_t(y) = y - s_t(y)$. Note that, even though the surplus function $s_t(y)$ is strictly increasing in y, the pricing function need not be strictly increasing or strictly decreasing in y.[7] Therefore, the price distribution $G_t(p)$ has to be written as an integral of the quality density $\Phi'(y)$ over the qualities y such that $p_t(y)$ is smaller than p.

2.3 Market dynamics

In this subsection, we characterize the dynamics of equilibrium. In particular, we are interested in the evolution of the distribution of surplus offered by sellers, the distribution of prices posted by sellers, the distribution of transactions across sellers and the evolution of sales concentration.

The dynamics of equilibrium are driven by the accumulation of the amount of information that buyers have about sellers in the market. In period t, a buyer is in contact with n_t sellers, where n_t is distributed as a Poisson with coefficient λ_t. In period $t+1$, a buyer is in contact with n_{t+1} sellers, where n_{t+1} is distributed as a Poisson with coefficient λ_{t+1}. As long as buyers have some memory of past contacts, in the sense that $\delta < 1$, the expression in Lemma 1 shows that λ_t is strictly increasing over time. Since the coefficient λ_t is equal to the buyer's average number of contacts in period t, the fact that λ_t is strictly increasing over time means that, on average, buyers are in contact with an increasing number of sellers.

As buyers accumulate contacts, the market becomes more competitive over time and, for this reason, the surplus offered by sellers increases and, for the same reason, the prices posted by sellers decline. To see this, it follows from (2.11) that

$$s_t'(y) = \Phi'(y)\lambda_t(y - s_t(y)), \tag{2.15}$$

and

$$s_{t+1}'(y) = \Phi'(y)\lambda_{t+1}(y - s_{t+1}(y)), \tag{2.16}$$

together with the boundary conditions $s_t(y_\ell) = s_{t+1}(y_\ell) = 0$. The surplus function $s_{t+1}(y)$ is strictly greater than $s_t(y)$ for every $y \in (y_\ell, y_h]$. To see why this is the case, the following two observations are sufficient. First, notice that $s_{t+1}(y) = s_t(y)$ implies $s_{t+1}'(y) > s_t'(y)$, since $\lambda_{t+1} > \lambda_t$. In turn, this implies that $s_{t+1}(y)$ can equal $s_t(y)$ at most once, say for $y = y_c$, and that $s_{t+1}(y) > s_t(y)$ for all $y > y_c$. Second, notice that $s_{t+1}(y_\ell) = s_t(y_\ell)$. Taken together, these two observations imply that $s_{t+1}(y) > s_t(y)$ for all $y \in (y_\ell, y_h]$.

[7] If the extent of heterogeneity in the sellers' quality is small enough, the pricing function $p_t(y)$ is strictly decreasing in y.

Since the surplus function $s_{t+1}(y)$ is strictly greater than $s_t(y)$, it follows from (2.13) that $F_{t+1}(s) < F_t(s)$. Since $s_{t+1}(y)$ is strictly greater than $s_t(y)$, the price function $p_{t+1}(y) = y - s_{t+1}(y)$ is strictly smaller than $p_t(y)$. It then follows from (2.14) that $G_{t+1}(p) > G_t(p)$.

Proposition 1: (Surplus and price dynamics with memory) *Let $\delta \in [0, 1)$. For any given $y > y_\ell$, the surplus $s_t(y)$ is strictly increasing in t and the price $p_t(y)$ is strictly decreasing in t. The surplus distribution $F_t(s)$ is strictly increasing in t in the sense of first-order stochastic dominance, and the price distribution $G_t(p)$ is strictly decreasing in t in the sense of first-order stochastic dominance.*

As buyers accumulate contacts, they are not only offered more surplus by each seller, but they also get to purchase the good from sellers that offer more surplus, which are those that carry a variety of the good of higher quality. Both the increase in the surplus offered by each seller of a given quality y and the shift of transactions from low to high-surplus sellers (or, equivalently, from low to high-quality sellers) contribute to increase the buyer's surplus.

Proposition 2: (Transaction dynamics with memory) *For any $\delta \in [0, 1)$, the distribution of transactions across sellers of different quality, $H_t(y)$, is strictly increasing in t in the sense of first-order stochastic dominance.*

As shown in Proposition 2, the distribution of transactions moves towards higher quality sellers as buyers accumulate contacts. Since high quality sellers trade more than low quality sellers, trade becomes more concentrated as buyers accumulate contacts. This finding may appear surprising, since the market becomes progressively more competitive. Indeed, it is true that, as buyers accumulate contacts and their choice sets expand, the increase in competition drives prices down. At the same time, though, as buyers accumulate contacts and their choice sets expand, buyers can purchase the good from better sellers. Since all buyers have the same ranking over sellers, trade concentration increases. Indeed, in a perfectly competitive version of the market, only the highest-quality seller would trade.

To formalize this intuition, the quantity of the good traded by a seller with a variety of quality y in period t is

$$q_t(y) = b\lambda_t e^{-\lambda_t(1-\Phi(y))}. \tag{2.17}$$

Clearly, the quantity $q_t(y)$ is strictly increasing in y because $\Phi'(y) > 0$. Since the quantity traded by a seller is a strictly increasing function of the seller's quality y, it follows that the largest sellers are the sellers with the highest quality. Hence, the fraction of $Q_t(x)$ of transactions made by the fraction x of largest sellers is given by

$$Q_t(x) = 1 - H_t(y(1-x)), \tag{2.18}$$

where $y(z) = \Phi^{-1}(z)$, the quality of a seller at the z-th quantile of the distribution.

In Proposition 2, we established that $H_{t+1}(y) < H_t(y)$ for all $y \in (y_\ell, y_h)$. Therefore, $Q_{t+1}(x) > Q_t(x)$ for all $x \in (0,1)$.

Proposition 3: (Concentration dynamics with memory) *For any $\delta \in [0,1)$, sales concentration is increasing over time, i.e. the fraction $Q_t(x)$ of sales made by the fraction x of the largest sellers is strictly increasing over time.*

We now want to examine the role of memory in shaping the equilibrium of the product market. Let us first compare the properties of equilibrium if buyers have some memory of the sellers that they have contacted in the past, in the sense that $\delta \in [0,1)$, with the properties of equilibrium if buyers have no memory, in the sense that $\delta = 1$.

In period 1, the buyers' number of contacts is distributed as a Poisson with coefficient $\lambda_1 = \mu$ whether buyers have memory or not. For this reason, in period 1, the surplus distribution, the price distribution and the transaction distribution are the same whether buyers have memory or not. Formally, $F_1(s|\delta)$, $G_1(p|\delta)$ and $H_1(y|\delta)$ are independent of δ. In periods $t \geq 2$, memory matters. With memory, i.e. $\delta \in [0,1)$, we have $\lambda_2 > \lambda_1$, $\lambda_3 > \lambda_2$, etc... Without memory, i.e. $\delta = 1$, we have $\lambda_2 = \lambda_1$, $\lambda_3 = \lambda_2$, etc... For this reason, with memory, the market becomes more competitive, the surplus offered by sellers increases, the prices posted by sellers decline, and buyers shift their transactions from lower to higher quality sellers. Formally, $F_t(s|\delta)$ is strictly decreasing in t, $G_t(p|\delta)$ is strictly increasing in t, and $H_t(y|\delta)$ is strictly decreasing in t. Without memory, i.e. $\delta = 1$, the surplus distribution $F_t(s|\delta)$ remains equal to $F_1(s|\delta)$, the price distribution $G_t(p|\delta)$ remains equal to $G_1(p|\delta)$, and the transaction distribution $H_t(y|\delta)$ remains equal to $H_1(y|\delta)$.

Next, let us compare the properties of equilibrium if buyers have better or worse memory. That is, let us compare the equilibrium if buyers have memory $1-\delta_1$ or $1-\delta_2$, with $\delta_1 < \delta_2$. In period 1, $F_1(s|\delta_1) = F_1(s|\delta_2)$, $G_1(p|\delta_1) = G_1(p|\delta_2)$ and $H_1(y|\delta_1) = H_1(y|\delta_2)$. In periods $t \geq 2$, λ_t is strictly greater if buyers have better memory (δ_1) than if buyers have worse memory (δ_2). Therefore, in any period $t \geq 2$, surplus is higher, prices are lower, and transactions are better if buyers have better memory (δ_1) than if they have worse memory (δ_2). Formally, for $t \geq 2$, $F_t(s|\delta_1) < F_t(s|\delta_2)$, $G_t(p|\delta_1) > G_t(p|\delta_2)$ and $H_t(y|\delta_1) < H_t(y|\delta_2)$.

The above observations are summarized in the following proposition.

Proposition 4: (Memory). *Take any δ_1 and δ_2, with $0 \leq \delta_1 < \delta_2 \leq 1$.*

(i) *In period 1, equilibrium surplus, prices and transactions are the same for δ_1 and δ_2, in the sense that $F_1(s|\delta_1) = F_1(s|\delta_2)$, $G_1(p|\delta_1) = G_1(p|\delta_2)$ and $H_1(y|\delta_1) = H_1(y|\delta_2)$.*

(ii) *In period $t \geq 2$, equilibrium surplus is higher, prices are lower, and transactions are better with better memory (i.e. lower δ), in the sense that $F_t(s|\delta_1) < F_t(s|\delta_2)$, $G_t(p|\delta_1) > G_t(p|\delta_2)$ and $H_t(y|\delta_1) < H_t(y|\delta_2)$.*

(iii) *Let* $\delta_2 = 1$. *Without memory, equilibrium surplus, prices and transactions are constant over time, in the sense that* $F_t(s|\delta_2) = F_1(s|\delta_2)$, $G_t(p|\delta_2) = G_1(p|\delta_2)$, *and* $H_t(y|\delta_2) = H_1(y|\delta_2)$.

2.4 Steady state

Lastly, we want to understand the limiting behavior of the product market. In particular, we want to understand whether the market eventually becomes perfectly competitive, in the sense that every buyer purchases the good from the seller with the highest quality at a price equal to the seller's marginal cost.

To understand the limiting behavior of the product market, it is sufficient to examine the limiting behavior of the average number of contacts, λ_t. If buyers have imperfect memory of sellers met in the past, in the sense that $\delta \in (0, 1)$, then λ_t is strictly increasing and converges to $\lambda^* = \mu/\delta$.[8] Given that $H_{t+1}(y) < H_t(y)$ for all t and $H_t(y) \in [0, 1]$, it is easy to see that $H_t(y)$ converges to a stationary distribution $H^*(y)$.

Proposition 5: (Ergodic transaction distribution with memory) *For any* $\delta > 0$, *the transaction distribution* $H_t(y)$ *converges to an ergodic distribution* $H^*(y)$ *that is given by*

$$H^*(y) = \frac{e^{-\lambda^*(1-\Phi(y))} - e^{-\lambda^*}}{1 - e^{-\lambda^*}} \tag{2.19}$$

for all $y \in [y_\ell, y_h]$. *The distribution* $H^*(y)$ *is atomless and* $\lambda^* = \mu/\delta$.

Since λ_t is strictly increasing, the surplus offered by sellers increases over time, the prices posted by sellers decline over time, and the distribution of transactions moves from lower to higher quality sellers. Given that $s_{t+1}(y) > s_t(y)$ for all t and $s_t(y) \in [0, y_h]$, the surplus function $s_t(y)$ converges to a stationary surplus function $s^*(y)$ given by (2.12) where $\lambda^* = \mu/\delta$. Since λ^* is finite, the equilibrium does not become perfectly competitive. Indeed, since λ^* is finite, a seller with any quality $y \in [y_\ell, y_h]$ meets a strictly positive measure $b\lambda^* \exp(-\lambda^*)$ of captive buyers and, hence, can guarantee itself a strictly positive profit. For this reason, the surplus function $s^*(y)$ is such that $s^*(y) < y$, the price function $p^*(y)$ is such that $p^*(y) > 0$, and the volume of sales, $b\lambda^* \exp(-\lambda^*(1 - \Phi(y)))$, is strictly positive, for all $y \in [y_\ell, y_h]$. Only in the frictionless limit as $\lambda^* \to \infty$ is the equilibrium perfectly competitive. As $\lambda^* \to \infty$, the distribution of transactions $H^*(y)$ is a degenerate distribution at y_h and $s^*(y_h) \to \tilde{y}^* = y_h$ and $p^*(y_h) \to 0$.

3 Word of mouth

In this section, we construct a version of the search-theoretic model of Butters (1977), Varian (1980) and Burdett and Judd (1983) in which buyers learn about sellers not only by directly searching the market but also by talking to buyers that were active in the market

[8] To see this, Lemma 1 implies that $\lambda^* = \lim_{t\to\infty} \lambda_t = \frac{\mu}{\delta} \lim_{t\to\infty} [1 - (1 - \delta)^t] = \frac{\mu}{\delta}$.

in the previous period. We construct the model so that, even though sellers are long-lived, the seller's problem of choosing what price to post remains static. We establish that the equilibrium exists and is unique, and we characterize the evolution of equilibrium outcomes over time. We show that memory and word of mouth generate similar dynamics along some dimensions (i.e., concentration), but not along other dimensions (i.e., competition and prices).

3.1 Environment

We consider the market for some consumer good. The market is populated by a continuum b of infinitely-lived sellers with measure 1. Sellers are heterogeneous with respect to the quality y of their variety of the good. The distribution of sellers with respect to y is given by a cumulative distribution function $\Phi(y)$ with $\Phi'(y) > 0$ and support $[y_\ell, y_h]$ where $0 < y_\ell < y_h$. A seller produces the good at a constant marginal cost, which we set to 0. In every period t, a seller posts a price p.[9] If a seller trades q units of the good at the price p, its periodical profit is qp.

The market is also populated by a double continuum of short-lived buyers with measure b per seller.[10] New buyers enter the market in period t. If one of these buyers purchases a variety of the good of quality y at the price p, its periodical utility is $y - p$. If one of these buyers does not purchase the good, its periodical utility is 0. Whether it purchases the good or not, the buyer exits the market at the end of period t.

The product market is frictional, in the sense that a buyer cannot purchase the good from any seller in the market, but only from those sellers with which it is in contact. In every period t, a buyer may come into contact with sellers through two channels: search and word of mouth. Through search, a buyer comes into contact with m_t sellers, where m_t is a random variable distributed as a Poisson with coefficient $\mu > 0$. In every period t, a buyer comes into contact with r_t additional sellers through word of mouth, where r_t is a random variable distributed as a Poisson with coefficient $\rho > 0$. Since there are no old buyers in period 1, the word of mouth channel is inactive and the r_t sellers contacted in period 1 are found through random search. The average number of contacts in period t through either channel is $\lambda = \mu + \rho$, where λ is constant over time.

Depending on whether a contact is obtained through either search or word of mouth, the sellers in buyers' choice sets are potentially drawn from different distributions. Each contact obtained through search is randomly selected from the seller's distribution of qualities $\Phi(y)$. Each contact obtained through word of mouth is randomly selected from the distribution $\hat{\Phi}_t(y)$, where $\hat{\Phi}_t(y)$ denotes the distribution of the highest quality seller

[9] We assume that sellers are not allowed to post price schedules. In particular, a seller is not allowed to post a price schedule that depends on the calendar time t, or on the purchasing history h_t of an individual buyer.

[10] We assume that there is a continuum of buyers per seller. The assumption guarantees that an individual seller cannot learn anything about the demand that it faces from the quantity of the good that it has sold in the past, since such quantity is a deterministic function of the seller's quality and price.

contacted by buyers who purchased in the previous period. We conjecture (and later verify) that $\hat{\Phi}_t(y)$ is twice differentiable. In period 1, note that $\hat{\Phi}_1(y) = \Phi(y)$.

The environment described above is quite natural. A buyer that is new to the market can find out about sellers by searching the market directly. The outcome of this direct search process is a number m of meetings with randomly-selected seller. This is the same search process as in Butters (1977), Varian (1980) or Burdett and Judd (1983). A buyer that is new to the market, however, can also find out about sellers indirectly by looking for buyers that successfully purchased in the market in the previous period. The outcome of this indirect search process is a number r of meetings with previous buyers. In the meeting, the new buyer learns about the highest quality seller with which the old buyer had been in contact in the previous period.[11]

We assume that an old buyer refers a new buyer to the one seller that, among its contacts, carries the variety of the good with the highest quality, rather than to the seller that, among its contacts, offered the highest surplus in the previous period. The assumption simplifies the analysis. In fact, if old buyers refer new buyers to the highest-quality seller of which they know, the seller's price in the current period does not affect the measure and type of buyers that reach the seller in the future. Therefore, if old buyers refer new buyers to the highest-quality seller of which they know, the seller's pricing problem is static. The assumption is also easy to rationalize as an equilibrium. If sellers do not expect the current price to affect their future demand, higher quality sellers always choose to offer more surplus to their buyers. In turn, if higher quality sellers are expected to offer more surplus to their buyers, new buyers ask old buyers about the highest-quality seller of which they know, since that seller is the one expected to offer the highest surplus to the new buyer.

3.2 Equilibrium

Consider some arbitrary period $t = 1, 2, \ldots$. The distribution of sellers across qualities is given by $\Phi(y)$. This is the distribution from which buyers sample when they contact a seller through the search channel. The distribution of the highest quality seller contacted by buyers who purchased in period $t-1$ is given by $\hat{\Phi}_t(y)$. This is the distribution from which buyers sample when they contact a seller through word of mouth.

We denote as $F_t(s)$ the distribution of surplus offered by sellers. For the same reasons as in Section 2, the distribution $F_t(s)$ does not have any mass points. We denote as $\hat{F}_t(s)$ the distribution of surplus offered by sellers reached through word of mouth, which also cannot have any mass points.

Consider a seller of quality y that offers the surplus $s \geq 0$ in some period $t = 2, 3 \ldots$. The seller is reached by buyers through both search and word of mouth. Let us first

[11] It seems quite natural that a new buyer would only ask an old buyer about the best seller of which it knows. Information about additional sellers does not improve the new buyer's payoffs.

consider the buyers that reach the seller through search. There is a measure $b_{k,t}$ of buyers that reach the seller through search and that contacted k additional sellers through search, where $b_{k,t}$ is given by

$$b_{k,t} = b(k+1)\frac{e^{-\mu}\mu^{k+1}}{(k+1)!}. \tag{3.1}$$

Let us explain the expression in (3.1). Per seller, there is a measure $b\exp(-\mu)\mu^{k+1}/(k+1)!$ of buyers that have $k+1$ contacts through search. Therefore, per seller, there is a measure $b(k+1)\exp(-\mu)\mu^{k+1}/(k+1)!$ of search contacts that are generated by buyers that have met $k+1$ sellers through search. Since every seller has the same probability of being contacted through search, the measure of buyers that reach a particular seller through search and that have met k additional sellers through the search channel is given by the expression in (3.1).

It is easy to compute the probability that one of the $b_{k,t}$ buyers purchases the good from the seller. A fraction $\exp(-\rho)$ of the $b_{k,t}$ buyers did not contact any sellers through word of mouth. These buyers purchase the good from the seller with probability $F_t(s)^k$, where $F_t(s)$ denotes the distribution of surplus offered by sellers. A fraction $\exp(-\rho)\rho$ of the $b_{k,t}$ buyers did contact one seller through word of mouth. These buyers purchase the good from the seller with probability $F_t(s)^k\hat{F}_t(s)$, where $\hat{F}_t(s)$ is the distribution of surplus offered by sellers contacted through word of mouth. More generally, a fraction $\exp(-\rho)\rho^j/j!$ of the $b_{k,t}$ buyers did contact j sellers through word of mouth, with $j = 2, 3, \ldots$. These buyers purchase the good from the seller with probability $F_t(s)^k\hat{F}_t(s)^j$. Overall, the probability that one of the $b_{k,t}$ buyers purchases the good from the seller is given by

$$\pi_{k,t}(s) = \sum\nolimits_{j=0}^{\infty} \frac{e^{-\rho}\rho^j}{j!} F_t(s)^k \hat{F}_t(s)^j. \tag{3.2}$$

Next, consider the buyers that reach the seller through word of mouth. There is a measure $\hat{b}_{k,t}$ buyers that reach the seller through word of mouth and that reached k other sellers through word of mouth, where $\hat{b}_{k,t}$ is given by

$$\hat{b}_{k,t} = b(k+1)\kappa_t(y)\frac{e^{-\rho}\rho^{k+1}}{(k+1)!} \tag{3.3}$$

and $\kappa_t(y)$ is given by

$$\kappa_t(y) = \frac{\hat{\Phi}_t'(y)}{\Phi'(y)}. \tag{3.4}$$

Let us explain the expression in (3.3). Per seller, there is a measure $b\exp(-\rho)\rho^{k+1}/(k+1)!$ of buyers that have $k+1$ contacts through word of mouth. Therefore, per seller, there is a measure $b(k+1)\exp(-\rho)\rho^{k+1}/(k+1)!$ of word-of-mouth contacts that are generated by buyers that are in contact with $k+1$ sellers through word of mouth. If every seller has the same probability of being contacted through word of mouth, the measure of buyers that reach a particular seller through word of mouth and that have contacted an additional k

sellers through the word-of-mouth channel would be $b(k+1)\exp(-\rho)\rho^{k+1}/(k+1)!$.

The term $\kappa_t(y)$ given by (3.4) reflects the fact that some sellers are more likely than others to be contacted through word of mouth. In particular, the fraction of word-of-mouth contacts that reaches sellers with quality $\hat{y} \in [y, y+\epsilon]$ is $\hat{\Phi}(y+\epsilon) - \hat{\Phi}(y)$, while the measure of sellers with quality $\hat{y} \in [y, y+\epsilon]$ is $\Phi(y+\epsilon) - \Phi(y)$. Therefore, the measure of buyers that reach a seller of quality $\hat{y} \in [y, y+\epsilon]$ through word of mouth and that contacted an additional k sellers through word of mouth is $b(k+1)\exp(-\rho)\rho^{k+1}/(k+1)!$ multiplied by $[\hat{\Phi}(y+\epsilon) - \hat{\Phi}(y)]/[\Phi(y+\epsilon) - \Phi(y)]$. Taking the limit for $\epsilon \to 0$ yields $\kappa_t(y) = \hat{\Phi}_t'(y)/\Phi'(y)$. The term $\kappa_t(y)$ will turn out to be important later on.

It is easy to compute the probability that one of the $\hat{b}_{k,t}$ buyers purchases the good from the seller. A fraction $\exp(-\mu)$ of the $\hat{b}_{k,t}$ buyers did not contact any sellers through search. These buyers purchase the good from the seller with probability $\hat{F}_t(s)^k$. A fraction $\exp(-\mu)\mu$ of the $\hat{b}_{k,t}$ buyers did contact one seller through search. These buyers purchase the good from the seller with probability $\hat{F}_t(s)^k F_t(s)$. More generally, a fraction $\exp(-\mu)\mu^j/j!$ of the $\hat{b}_{k,t}$ buyers did contact j sellers through search. These buyers purchase the good from the seller with probability $\hat{F}_t(s)^k F_t(s)^j$. Overall, the probability that one of the $\hat{b}_{k,t}$ buyers purchases the good from the seller is given by

$$\hat{\pi}_{k,t}(s) = \sum\nolimits_{j=0}^{\infty} \frac{e^{-\mu}\mu^j}{j!} F_t(s)^j \hat{F}_t(s)^k. \tag{3.5}$$

We can now write the profit $V_t(y,s)$ for a seller with quality y offering the surplus $s \geq 0$ is given by

$$\begin{aligned} V_t(y,s) &= \left[\textstyle\sum_{k=0}^{\infty} b_{k,t}\pi_{k,t}(s)\right](y-s) + \left[\textstyle\sum_{k=0}^{\infty} \hat{b}_{k,t}\hat{\pi}_{k,t}(s)\right](y-s) \\ &= \left[\textstyle\sum_{k=0}^{\infty} b(k+1)\dfrac{e^{-\mu}\mu^{k+1}}{(k+1)!}\left(\textstyle\sum_{j=0}^{\infty}\dfrac{e^{-\rho}\rho^j}{j!}F_t(s)^k\hat{F}_t(s)^j\right)\right](y-s) \\ &+ \left[\textstyle\sum_{k=0}^{\infty} b(k+1)\kappa_t(y)\dfrac{e^{-\rho}\rho^{k+1}}{(k+1)!}\left(\textstyle\sum_{j=0}^{\infty}\dfrac{e^{-\mu}\mu^j}{j!}F_t(s)^j\hat{F}_t(s)^k\right)\right](y-s). \end{aligned} \tag{3.6}$$

Collecting terms, we can rewrite (3.6) as

$$\begin{aligned} V_t(y,s) &= \left[\textstyle\sum_{k=0}^{\infty} b\mu e^{-\mu}\dfrac{\mu^k F_t(s)^k}{k!}e^{-\rho(1-\hat{F}_t(s))}\left(\textstyle\sum_{j=0}^{\infty}\dfrac{e^{-\rho\hat{F}_t(s)}\rho^j\hat{F}_t(s)^j}{j!}\right)\right](y-s) \\ &+ \left[\textstyle\sum_{k=0}^{\infty} b\rho\kappa_t(y)e^{-\rho}\dfrac{\rho^k\hat{F}_t(s)^k}{k!}e^{-\mu(1-F_t(s))}\left(\textstyle\sum_{j=0}^{\infty}\dfrac{e^{-\mu F_t(s)}\mu^j F_t(s)^j}{j!}\right)\right](y-s) \\ &= \left[\textstyle\sum_{k=0}^{\infty} b\mu e^{-\mu}\dfrac{\mu^k F_t(s)^k}{k!}e^{-\rho(1-\hat{F}_t(s))}\right](y-s) \\ &+ \left[\textstyle\sum_{k=0}^{\infty} b\rho\kappa_t(y)e^{-\rho}\dfrac{\rho^k\hat{F}_t(s)^k}{k!}e^{-\mu(1-F_t(s))}\right](y-s), \end{aligned} \tag{3.7}$$

where the second equality makes use of the fact that the summations over j in lines 1 and

3 are both equal to 1. Collecting terms once more, we can rewrite (3.7) as

$$\begin{aligned} V_t(y,s) &= b\mu e^{-\mu(1-F_t(s))}e^{-\rho(1-\hat{F}_t(s))}\left[\sum_{k=0}^{\infty} e^{-\mu F_t(s)}\frac{\mu^k F_t(s)^k}{k!}\right](y-s) \\ &\quad +b\rho\kappa_t(y)e^{-\mu(1-F_t(s))}e^{-\rho(1-\hat{F}_t(s))}\left[\sum_{k=0}^{\infty} e^{-\rho\hat{F}_t(s)}\frac{\rho^k \hat{F}_t(s)^k}{k!}\right](y-s) \\ &= b\left(\mu+\rho\kappa_t(y)\right)e^{-\mu(1-F_t(s))}e^{-\rho(1-\hat{F}_t(s))}(y-s). \end{aligned} \tag{3.8}$$

where the second equality makes use of the fact that the summations over k in the second and third lines are both equal to 1.

Using the same arguments as in Lemma 4, it is straightforward to show that the support of the surplus distribution $F_t(s)$ is some interval $[s_{\ell,t}, s_{h,t}]$, with $s_{\ell,t} = 0$. In turn, using the same arguments as in Lemma 5, it is easy to show that the surplus offered by a seller is a strictly increasing function $s_t(y)$ of the quality of the seller's variety y. From the strict monotonicity of $s_t(y)$, it follows that the fraction of sellers offering a surplus smaller than $s_t(y)$ is equal to the fraction of sellers with quality smaller than y, i.e.

$$F_t(s_t(y)) = \Phi(y). \tag{3.9}$$

Similarly, in the distribution of sellers reached through word of mouth, the fraction offering a surplus smaller than $s_t(y)$ is equal to the fraction of sellers reached through word of mouth with quality smaller than y, i.e.

$$\hat{F}_t(s_t(y)) = \hat{\Phi}_t(y). \tag{3.10}$$

Differentiating both (3.9) and (3.10) with respect to y yields

$$F_t'(s_t(y)) = \frac{\Phi'(y)}{s_t'(y)}, \quad \hat{F}_t'(s_t(y)) = \frac{\hat{\Phi}'(y)}{s_t'(y)}. \tag{3.11}$$

For a seller producing quality y, the seller's problem at date t is

$$V_t(y) = \max_{s\in[0,y]} b\left(\mu+\rho\kappa_t(y)\right)e^{-\mu(1-F_t(s))}e^{-\rho(1-\hat{F}_t(s))}(y-s). \tag{3.12}$$

The necessary condition for the optimality of the surplus $s_t(y)$ offered by a seller with quality y is

$$\begin{aligned} &\left[\mu F_t'(s_t(y)) + \rho\hat{F}_t'(s_t(y))\right] b\left(\mu+\rho\kappa_t(y)\right)e^{-\mu(1-F_t(s))}e^{-\rho(1-\hat{F}_t(s))}(y-s_t(y)) \\ = \;& b\left(\mu+\rho\kappa_t(y)\right)e^{-\mu(1-F_t(s))}e^{-\rho(1-\hat{F}_t(s))}. \end{aligned} \tag{3.13}$$

The left-hand side of (3.13) is the benefit to the seller from offering an extra unit of surplus to the buyers, which is given by the increase in the seller's volume multiplied by the seller's profit per trade. The right-hand side of (3.13) is the cost to the seller from offering an extra unit of surplus, which is given by the seller's volume. Using (3.9), (3.10)

and (3.11), we can rewrite (3.13) as

$$s_t'(y) = \Phi'(y)(\mu + \rho\kappa_t(y))(y - s_t(y)). \tag{3.14}$$

Since buyers purchase from the contacted seller that offers the highest surplus, and sellers with higher quality offer higher surplus, a buyer purchases the good from the contacted seller that carries the variety of the good with highest quality. Since the transaction distribution is equal to the distribution of the highest quality seller contacted by a buyer in the previous period, conditional on the buyer having contacted at least one seller, the sampling distribution $\hat{\Phi}_t(y)$ for sellers contacted in periods $t > 1$ through word of mouth is just the distribution of transactions in the previous period, $H_{t-1}(y)$. In period 1, recall that $\hat{\Phi}_1(y) = \Phi(y)$ and therefore we use the convention that $H_0(y) = \Phi(y)$ in what follows.

Lemma 6: *The distribution $H_t(y)$ of transactions at time $t \geq 1$ is atomless and it is given by*

$$H_t(y) = \frac{e^{-\mu(1-\Phi(y))}e^{-\rho(1-H_{t-1}(y))} - e^{-(\mu+\rho)}}{1 - e^{-(\mu+\rho)}}. \tag{3.15}$$

The analysis above leaves us with a unique candidate equilibrium in any period $t \geq 1$. In the candidate equilibrium, the transaction distribution $H_t(y)$ is given by (3.15) and the surplus function $s_t(y)$ is given by the solution to the differential equation

$$s_t'(y) = \Phi'(y)(\mu + \rho\kappa_t(y))(y - s_t(y)), \tag{3.16}$$

where $H_0(y) = \Phi(y)$ and

$$\kappa_t(y) = \frac{H_{t-1}'(y)}{\Phi'(y)}, \tag{3.17}$$

together with the boundary condition $s_t(y_\ell) = 0$. As we did in Section 2, it is easy to verify that the unique candidate equilibrium is, indeed, an equilibrium since (3.16) is not only necessary but also sufficient for profit maximization. Note that in period 1, we have $\kappa_1(y) = 1$ for all $y \in [y_\ell, y_h]$ because we use the convention that $H_0(y) = \Phi(y)$.

Before we present Theorem 2, it is useful to first define $\beta = \rho/\lambda$, the average proportion of contacts reached through word of mouth, and $1 - \beta = \mu/\lambda$, the average proportion of contacts reached through search.

Theorem 2: (Existence and uniqueness of equilibrium with word of mouth) *An equilibrium exists and is unique. The equilibrium is such that:*

(i) *In period $t \geq 1$, we have $\hat{\Phi}_t(y) = H_{t-1}(y)$ and the surplus function $s_t(y)$ is the solution to the differential equation (3.16) that satisfies the boundary condition $s_t(y_\ell) = 0$.*

(ii) *Letting $\lambda = \mu + \rho$, the solution to the differential equation (3.16) is*

$$s_t(y) = \frac{\int_{y_\ell}^{y} \lambda e^{-\lambda(1-\Gamma_t(x))} x\Gamma_t'(x)dx}{e^{-\lambda(1-\Gamma_t(y))}} \tag{3.18}$$

where $\beta = \rho/\lambda$ *and*

$$\Gamma_t(y) = \beta H_{t-1}(y) + (1-\beta)\Phi(y), \tag{3.19}$$

where $s_t(y_h) = (1-e^{-\lambda})\tilde{y}_t$ *where* $\tilde{y}_t$ *is the average transaction quality.*

The solution $s_t(y)$ to the differential equation has the same structure as the solution in Theorem 1. Again, firms set prices $p_t(y) = y - s_t(y)$ equal to the expected value of the difference between their own quality y and the next best quality among the sellers a buyer contacts. The main difference is that the relevant distribution from which sellers are drawn is now $\Gamma_t(y)$, which is a weighted average of the distribution $H_{t-1}(y)$ of transactions in period $t-1$ and the distribution $\Phi(y)$ of seller qualities. We call $\Gamma_t(y)$ the *sampling distribution* for buyers. The weight β on the distribution $H_{t-1}(y)$ is the average proportion ρ/λ of contacts reached through word of mouth, while the weight $1-\beta$ on the distribution $\Phi(y)$ is the proportion of contacts obtained through search.

Associated with the equilibrium surplus function $s_t(y)$, there are a surplus distribution $F_t(s)$, a price distribution $G_t(p)$, and a transaction distribution $H_t(y)$. The surplus distribution $F_t(s)$ is given by

$$F_t(s) = \Phi(y_t(s)), \tag{3.20}$$

where $y_t(s)$ is the inverse of the surplus function $s_t(y)$. The price distribution $G_t(p)$ is given by

$$G_t(p) = \int_{y:p_t(y)\leq p} \Phi'(y)dy, \tag{3.21}$$

where $p_t(y)$ is the pricing function $y - s_t(y)$. The transaction distribution $H_t(y)$ can be expressed as

$$H_t(y) = \frac{e^{-\lambda(1-\Gamma_t(y))} - e^{-\lambda}}{1-e^{-\lambda}}. \tag{3.22}$$

For any period $t \geq 1$, the transaction distribution $H_t(y)$ first-order stochastically dominates the sampling distribution $\Gamma_t(y)$. For any period $t > 1$, the sampling distribution $\Gamma_t(y)$ also first-order stochastically dominates the distribution of seller qualities $\Phi(y)$ whenever $\rho > 0$. That is, for any period $t > 1$, we have $H_t(y) < \Gamma_t(y) < \Phi(y)$ for any $y \in (y_\ell, y_h)$ whenever word of mouth is active.

3.3 Market dynamics

In contrast to the model with memory, where the evolution of the average number of contacts λ_t drives market dynamics, it is the evolution of the sampling distribution $\Gamma_t(y)$ over time that drives market dynamics with word of mouth. Given that the distribution of seller qualities sampled through search is static, the evolution of the sampling distribution $\Gamma_t(y)$ is driven entirely by the dynamics of the transaction distribution $H_t(y)$, which determines the sellers sampled through word of mouth.

We have the following proposition regarding the evolution of $H_t(y)$ over time.

Proposition 6: (Transaction dynamics with word of mouth) *For any $\rho > 0$, the distribution of transactions across sellers of different quality, $H_t(y)$, is increasing over time in the sense of first-order stochastic dominance, and the sampling distribution $\Gamma_t(y)$ is also increasing over time in the sense of first-order stochastic dominance.*

Notice that the transaction distribution improves over time in the model with word of mouth just as it does in the model with memory. The reason why the transaction distribution improves over time is, however, different in the model with word of mouth and in the model with memory. In the model with memory, the transaction distribution improves over time because the buyers' choice sets become larger and larger and, for this reason, buyers can be more selective and purchase the good from better and better sellers. In the model with word of mouth, the transaction distribution improves over time not because the buyers' choice sets become larger and larger, but because the buyers' choice sets include better and better sellers. Through word of mouth, buyers sample the highest-quality sellers contacted by buyers in the previous period and they sample, through direct search, some additional sellers. For this reason, the highest-quality sellers sampled in period t are better than the highest-quality sellers sampled in period $t-1$.

Next, we examine the dynamics of trade concentration. In any period $t \geq 1$, the quantity of the good traded by a seller with a variety of quality y is

$$q_t(y) = b\left(\mu + \rho\kappa_t(y)\right) e^{-\mu(1-\Phi(y))} e^{-\rho(1-H_{t-1}(y))}. \tag{3.23}$$

In period 1, the quantity $q_t(y)$ is strictly increasing in y because $\Phi'(y) > 0$. For any period $t \geq 2$, the quantity $q_t(y)$ is strictly increasing in y because the term $\kappa_t(y)$ is strictly increasing in y. To see why this is the case, notice that (3.17) and (3.15) imply

$$\kappa_2(y) = \frac{H_1'(y)}{\Phi'(y)} = \frac{(\mu+\rho)e^{-\mu(1-\Phi(y))}e^{-\rho(1-H_1(y))}}{1-e^{-(\mu+\rho)}}. \tag{3.24}$$

It is evident from (3.24) that $\kappa_2(y)$ is strictly increasing in y. For any $t > 2$,

$$\kappa_{t+1}(y) = \frac{H_t'(y)}{\Phi'(y)} = \frac{(\mu+\rho\kappa_t(y))e^{-\mu(1-\Phi(y))}e^{-\rho(1-H_t(y))}}{1-e^{-(\mu+\rho)}}. \tag{3.25}$$

If $\kappa_t(y)$ is strictly increasing in y, it is clear from (3.25) that $\kappa_{t+1}(y)$ is also strictly increasing in y. By induction, we have that $\kappa_t(y)$ is strictly increasing in y for any $t \geq 2$. In period 1, we have $\kappa_t(y) = 1$ for all y.

Since the quantity of the good traded by a seller is a strictly increasing function $q_t(y)$ of the seller's quality y, it follows that the largest sellers are the sellers with the highest quality. Hence, the fraction of $Q_t(x)$ of transactions made by the x fraction of largest sellers is given by

$$Q_t(x) = 1 - H_t(\Phi^{-1}(1-x)), \tag{3.26}$$

where $y(z) = \Phi^{-1}(z)$, the quality of a seller at the z-th quantile of the distribution. Since

$H_t(y)$ is strictly deceasing in t, (3.26) implies that $Q_t(x)$ is strictly increasing in t.

Proposition 7: (Concentration dynamics with word of mouth) *For any $\rho > 0$, sales concentration is increasing over time, i.e. the fraction $Q_t(x)$ of sales made by the x fraction of largest sellers is strictly increasing over time.*

As in the model with memory, trade becomes more and more concentrated over time at higher quality sellers. In the model with memory, trade becomes more concentrated over time because the choice sets of buyers become larger and, hence, buyers can purchase the good from higher quality sellers, which are the sellers that offer more surplus and have higher trade volume. In the model with word of mouth, trade becomes more concentrated over time because the sellers in the choice sets of buyers have higher quality, and higher quality sellers are those that offer more surplus and have higher trade volume.

We now examine the dynamics of the surplus offered by sellers. In any period $t \geq 1$, the surplus function $s_t(y)$ is the solution to the differential equation

$$s_t'(y) = \Phi'(y)\left(\mu + \rho\kappa_t(y)\right)\left(y - s_t(y)\right) \tag{3.27}$$

and the surplus function $s_{t+1}(y)$ is the solution to

$$s_{t+1}'(y) = \Phi'(y)\left(\mu + \rho\kappa_{t+1}(y)\right)\left(y - s_t(y)\right), \tag{3.28}$$

where both satisfy the boundary conditions $s_t(y_\ell) = 0$. Since the integral of $\kappa_\tau(y) \equiv H_{\tau-1}'(y)/\Phi'(y)$ with respect to the seller's distribution $\Phi(y)$ is equal to 1 for all $\tau = 1, 2, ...,$ we know that $\kappa_{t+1}(y)$ cannot be greater than $\kappa_t(y)$ for all y. For this reason, the surplus function $s_{t+1}(y)$ need not be strictly greater than the surplus function $s_t(y)$ for all y.

Recall that $\kappa_t(y) = \hat{\Phi}_t'(y)/\Phi'(y)$, the ratio of the density of sellers contacted through word of mouth to the density of sellers contacted through search. It is useful to interpret $\kappa_t(y)$ as the *popularity* of a seller with quality y in period t because it is a measure of the extent to which the seller is more likely to be contacted through word of mouth than through search. The average popularity of a seller at time t is always equal to one. For sellers with higher qualities y that are more likely to be contacted through word of mouth, the density $\hat{\Phi}_t'(y)$ is greater than the density $\Phi'(y)$ and thus $\kappa_t(y) > 1$, i.e. their popularity is above average. For sellers with lower qualities y that are less likely to be contacted through word of mouth, the density $\hat{\Phi}_t'(y)$ is less than the density $\Phi'(y)$ and thus $\kappa_t(y) < 1$.

The dynamics of competition depend on whether a seller is increasing or decreasing in popularity $\kappa_t(y)$. Lemma 7 tells us that the highest quality sellers experience increasing popularity over time, while the lowest quality sellers experience decreasing popularity over time. For sellers of intermediate quality, their popularity varies non-monotonically over time, at first increasing and then eventually decreasing.

Lemma 7: (Dynamics of popularity with word of mouth) *For any $\rho > 0$, the pop-*

ularity of a seller $\kappa_t(y)$ may be increasing, decreasing, or non-monotonic over time. In particular, we have the following:

(i) *For any period t, there exists a unique cut-off quality $y_t^c \in (y_l, y_h)$ such that a seller's popularity is increasing at time t, i.e. $\kappa_{t+1}(y) > \kappa_t(y)$, if and only if $y > y_t^c$. The cut-off quality y_t^c is increasing over time.*

(ii) *If $y \geq y_{\max}^c = \lim_{t\to\infty} y_t^c$ then $\kappa_t(y)$ is always strictly increasing over time.*

(iii) *If $y < y_{\min}^c$ then $\kappa_t(y)$ is always strictly decreasing.*

(iv) *If $y \in (y_{\min}^c, y_{\max}^c)$ then $\kappa_t(y)$ is strictly increasing and then strictly decreasing.*

Proposition 8 says that the dynamics of competition are heterogeneous across sellers. Competition may be increasing or decreasing over time, or it may vary non-monotonically over time. At any given period of time, there is a range of low-quality sellers that experience decreasing competition and post a lower surplus and higher price over time, but there is also a range of high-quality sellers that experience increasing competition and post a higher surplus and a lower price over time. For the very lowest-quality sellers, competition is always decreasing over time, but for the highest-quality seller it is always increasing over time. Figure 1 illustrates an example of the heterogeneous price dynamics.

Proposition 8: (Dynamics of competition with word of mouth) *For any $\rho > 0$, the surplus function $s_t(y)$ and the price function $p_t(y)$ may be either increasing, decreasing, or non-monotonic over time. In particular, we have the following:*

(i) *For any period t, there exists a unique cut-off quality $y_t^s \in (y_l, y_h)$ such that, if $y > y_t^s$, the surplus $s_t(y)$ is strictly increasing at time t, i.e. $s_{t+1}(y) > s_t(y)$, and the price $p_t(y)$ is strictly decreasing at time t, i.e. $p_{t+1}(y) < p_t(y)$.*

(ii) *For any period t, if a seller's quality $y < y_t^c$, the surplus $s_t(y)$ is strictly decreasing and the price $p_t(y)$ is strictly increasing at time t.*

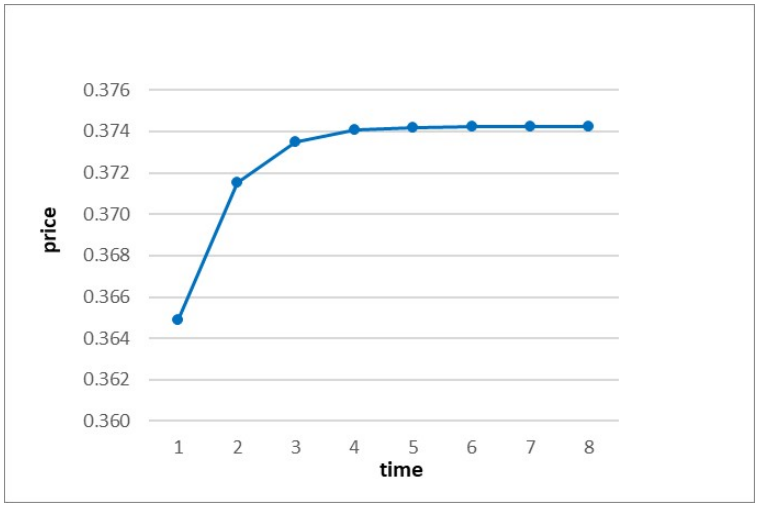
(iii) *If $y = y_h$, the surplus $s_t(y)$ is always strictly increasing and the price $p_t(y)$ is always strictly decreasing over time.*

(iv) *If $y < y_{\min}^c$, the surplus $s_t(y)$ is always strictly decreasing and the price $p_t(y)$ is always strictly increasing over time.*

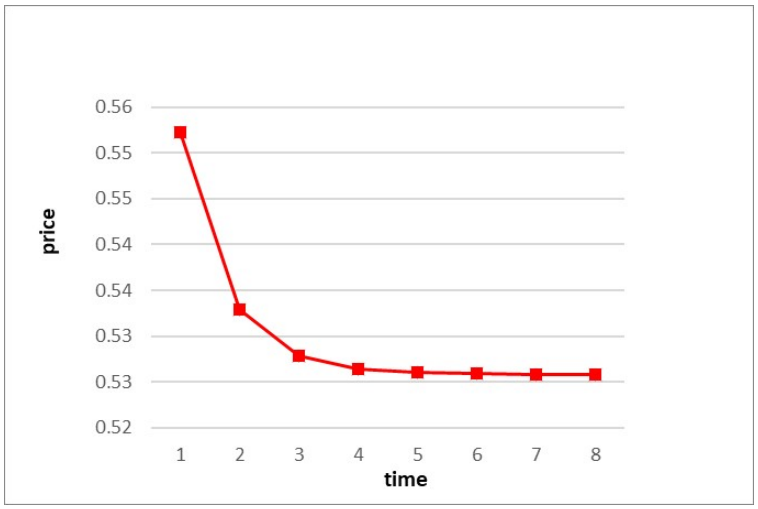
(v) *If $y \in (y_{min}^s, y_{max}^c)$, the surplus $s_t(y)$ and the price $p_t(y)$ are non-monotonic over time. In particular, the surplus (price) is increasing (decreasing) at first and then is eventually decreasing (increasing) over time.*

Why do more popular (higher quality) sellers tend to experience greater competition over time relative to less popular (lower quality) sellers? The answer lies not in a difference in the size of buyers' choice sets, but in the fact that word of mouth is a less competitive channel than search for sellers that are unpopular. To see this, consider what the surplus

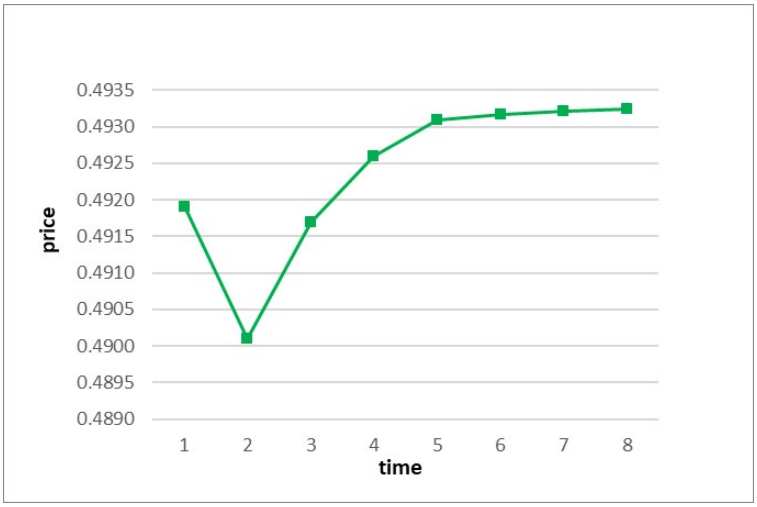
Figure 1: Price dynamics for word of mouth with $\rho = 1/3$, $\mu = 1$, uniform $\Phi(y)$ on $[0, 1]$



(a) Increasing prices for lower quality seller, $y = 0.5$



(b) Decreasing prices for highest quality seller, $y = 1$



(c) Non-monotonic prices for intermediate quality seller, $y = 0.8$

function would be if there was only the search channel and no word-of-mouth channel:

$$s_t'(y) = \Phi'(y)\mu\left(y - s_t(y)\right). \tag{3.29}$$

Now consider what the surplus function would be if there was no search channel and only the word-of-mouth channel:

$$s_t'(y) = H_{t-1}'(y)\rho\left(y - s_t(y)\right). \tag{3.30}$$

Suppose the average size of the choice set is the same, i.e. $\mu = \rho$. We know that $H_{t-1}'(y) < \Phi'(y)$ if and only if $\kappa_t(y) < 1$. It can be shown that, for unpopular sellers ($\kappa_t(y) < 1$), the surplus function $s_t(y)$ is strictly lower with word of mouth only versus search only, but for popular sellers ($\kappa_t(y) > 1$), $s_t(y)$ is strictly higher with word of mouth only versus search only.[12] For the best seller, the surplus function $s_t(y_h)$ is always higher at any time t with word of mouth only compared to search only.

In this sense, word of mouth is a more competitive channel than search for sellers that are popular. As a result, popular sellers, who are more likely to be contacted through word of mouth, anticipate steeper competition and offer buyers a higher surplus. The reason why there is greater competition and higher surplus offered by more popular sellers is the fact that the density of nearby sellers being sampled through word of mouth is higher for these sellers. Intuitively, this means that, with word of mouth, competing sellers are more likely to be close competitors with qualities y close to each other. This closeness between competitors reduces the gap between the highest and second-highest seller in a buyer's choice set, which reduces the price posted and increases the buyer's surplus.

Proposition 9: (Surplus and price dynamics with word of mouth). *For any $\rho > 0$, the surplus distribution $F_t(s)$ need not be increasing over time in the sense of first-order stochastic dominance, and the price distribution $G_t(p)$ need not be decreasing over time in the sense of first-order stochastic dominance.*

The above proposition highlights a critical difference between the model with memory and the model with word of mouth. In the model with memory, the choice set of buyers expands over time. The expansion of the buyers' choice sets increases the extent of competition and, in turn, leads to a decline in prices. Moreover, the expansion of the buyers' choice sets allows buyers to purchase from higher quality sellers and, in turn, leads to higher concentration. In the model with word of mouth, the choice set of buyers does not expand but it does improve, in the sense that it replaces lower-quality sellers with higher-quality sellers. The improvement in the buyers' choice sets directly leads to an increase in concentration. However, the improvement in the buyers' choice set affects sellers differently depending on their popularity. It does not increase competition for *all* sellers and, for this reason, does not necessarily lead to a decline in prices for all sellers. The highest quality sellers face steeper competition and post a higher surplus and lower

[12] See the Appendix for a proof.

price over time, while the lowest quality sellers face weaker competition and post a lower surplus and higher price over time.

It is useful to dive deeper in the comparison between the effect of memory and the effect of word of mouth on both market concentration and competition. Consider the model with memory. In equilibrium, the fraction x of the largest sellers trades a fraction $Q_t(x)$ of the good, where $Q_t(x)$ is given by

$$Q_t(x) = \frac{1 - e^{-\lambda_t x}}{1 - e^{-\lambda_t}}. \tag{3.31}$$

For x small, $Q_t(x)$ is approximately equal to

$$\begin{aligned} Q_t(x) &\simeq Q_t(0) - Q_t'(0)(0 - x) \\ &= (\lambda_t x)/(1 - e^{-\lambda_t}). \end{aligned} \tag{3.32}$$

The first line in (3.32) is a first-order approximation of $Q_t(x)$ around 0. The second line in (3.32) is obtained by noting that $Q_t(0) = 0$ and by computing the derivative of $Q_t(x)$ with respect to x and evaluating it at $x = 0$.

Now, consider the model with word of mouth. In equilibrium, the fraction x of the largest sellers trades a fraction $Q_t(x)$ of the good, where $Q_t(x)$ is given by

$$Q_t(x) = \frac{1 - e^{-\mu x} e^{-\rho(1 - H_{t-1}(y(1-x)))}}{1 - e^{-(\mu+\rho)}} \tag{3.33}$$

where $y(z) = \Phi^{-1}(z)$, the quality of a seller at the z-th quantile of the distribution. For x small, $Q_t(x)$ is approximately equal to

$$\begin{aligned} Q_t(x) &\simeq Q_t(0) - Q_t'(0)(0 - x) \\ &= \left(\mu + \rho(H_{t-1}'(y_h)/\Phi'(y_h))\right) x/(1 - e^{-(\mu+\rho)}) \\ &= (\mu + \rho\kappa_t(y_h))\, x/(1 - e^{-(\mu+\rho)}). \end{aligned} \tag{3.34}$$

The first line in (3.34) is a first-order approximation of $Q_t(x)$ around 0. The second line in (3.34) is obtained by noticing that $Q_t(0) = 0$ and by computing the derivative of $Q_t(x)$ and evaluating it at $x = 0$. The third line in (3.34) uses (3.17).

Imagine a situation in which the concentration of trade at the fraction x of the top sellers is the same in the equilibrium of the model with memory as in the equilibrium of the model with word of mouth, for $x \simeq 0$. From (3.32) and (3.34), it follows that the concentration of sales is the same in the two models if and only if

$$\frac{\lambda_t}{1 - e^{-\lambda_t}} = \frac{\mu + \rho\kappa_t(y_h)}{1 - e^{-(\mu+\rho)}}. \tag{3.35}$$

We know that $\kappa_t(y_h) > 1$ because $\kappa_t(y)$ is strictly increasing in y and the integral of $\kappa_\tau(y) \equiv H_{\tau-1}'(y)/\Phi'(y)$ with respect to the seller's distribution $\Phi(y)$ is equal to 1. If $\lambda_t \leq \mu + \rho$, then (3.35) cannot hold because $z/(1 - e^{-z})$ is increasing in z and $\kappa_t(y_h) > 1$.

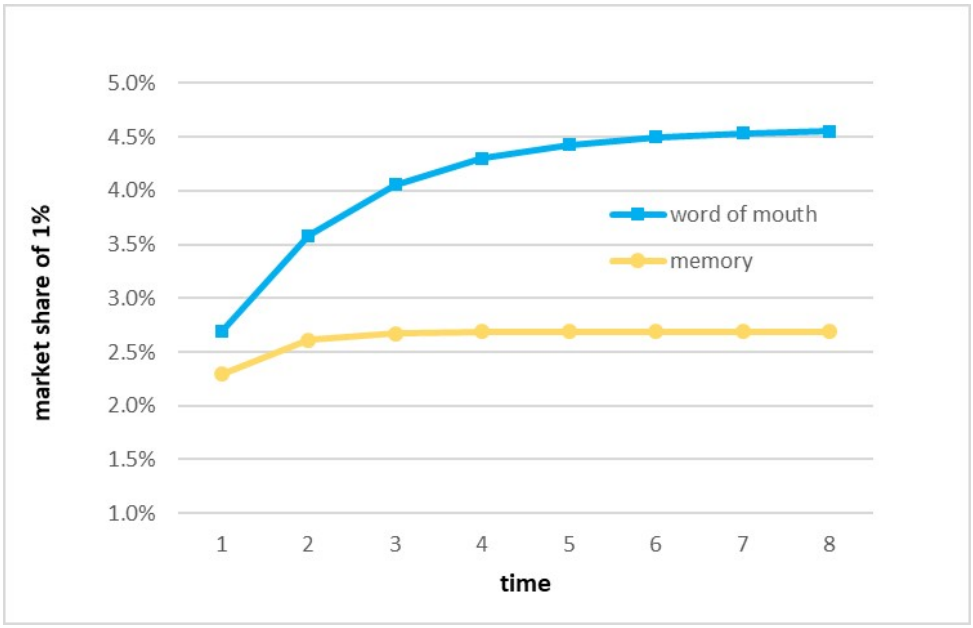


Figure 2: Market share of top 1% of sellers with $\mu = 2$, $\beta = 1 - \delta$, $\Phi(y)$ uniform

Therefore, we require $\lambda_t > \mu + \rho$ in order for (3.35) to hold. In fact, because this also implies $1 - e^{-\lambda_t} > 1 - e^{-(\mu+\rho)}$, condition (3.35) holds only if

$$\lambda_t > \mu + \rho\kappa_t(y_h). \tag{3.36}$$

This means that in order for the sales concentration to be the same across the two models, the model with memory requires a strictly greater average number of contacts λ_t.

In the model with memory, the average number of contacts λ_t depends on the memory parameter $1 - \delta$. In particular, we know that $\lambda_t = \frac{\mu}{\delta}[1 - (1 - \delta)^t]$ from Lemma 1. If the average number of new sellers reached through search, μ, is the same for both models, then condition (3.36) requires that

$$\frac{\mu}{\delta}[1 - (1 - \delta)^t] > \mu + \rho\kappa_t(y_h). \tag{3.37}$$

The left-hand side of (3.37) is less than μ/δ whenever $\delta > 0$. The right-hand side of (3.37) is greater than $\mu + \rho$ because $\kappa_t(y_h) > 1$. Therefore, if $\mu/\delta < \mu + \rho$, condition (3.36) cannot hold. This means that if $\beta > 1 - \delta$, i.e. the proportion of contacts β reached through word of mouth is greater than the memory parameter $1 - \delta$, we cannot have the same concentration for both models. In this case, concentration is always higher in the model with word of mouth compared to the model with memory. We summarize this observation in Proposition 10 below.

Figure 2 illustrates an example for the top 1% of sellers by quality where $\beta = 1 - \delta$.

Proposition 10: *Suppose the average number of sellers reached through search, μ, is the same for both memory and word of mouth. Sales concentration is always higher with*

word of mouth compared to memory if $\beta > 1 - \delta$.

If the proportion of contacts reached through word of mouth is sufficiently low, i.e. $\beta < 1-\delta$, the concentration of sales will be identical in the equilibrium of the two models (for small x) if (3.35) holds. However, the equilibrium distribution of surplus and prices is not the same. Indeed, in the equilibrium of the model with memory, the surplus function is given by the solution to the differential equation

$$s_t'(y) \quad = \Phi'(y)\lambda_t(y - s_t(y)), \tag{3.38}$$

with the boundary condition $s_t(y_\ell) = 0$, where $\lambda_t > \mu + \rho\kappa_t(y_h)$. In the equilibrium of the model with word of mouth, the surplus function is the solution to the differential equation

$$s_t'(y) = \Phi'(y)\left(\mu + \rho\kappa_t(y)\right)(y - s_t(y)), \tag{3.39}$$

with the boundary condition $s_t(y_\ell) = 0$. Since $\kappa_t(y)$ is strictly increasing in y, we have $\kappa_t(y_h) > \kappa_t(y)$ for all $y < y_h$ and it follows from (3.38) and (3.39) that the surplus function $s_t(y)$ is strictly greater in the model with memory than in the model with word of mouth. In turn, this implies that the surplus offered by sellers, for any given y, is strictly greater in the model with memory than in the model with word of mouth. That is, $F_t(s)$ is strictly smaller in the model with memory than in the model with word of mouth. Similarly, the prices posted by sellers are strictly lower, for any given y, in the model with memory than in the model with word of mouth. That is, $G_t(p)$ is strictly greater in the model with memory than in the model with word of mouth.

Proposition 11. *If sales concentration for a small fraction x of the top sellers is the same for both memory and word of mouth, then competition is strictly greater with memory compared to word of mouth. In particular, surplus offers $s_t(y)$ are higher and prices posted $p_t(y)$ are lower for all y, the surplus distribution $F_t(s)$ is higher in the sense of first-order stochastic dominance, and the price distribution $G_t(p)$ is lower in the sense of first-order stochastic dominance, with memory compared to word of mouth.*

The findings above reveal that, when word of mouth and memory generate the same degree of market concentration, word of mouth generates less competition than memory. The findings are intuitive. Concentration depends on the distribution of the best seller in the buyers' choice set. Competition depends not only on the best seller in the buyers' choice set, but also on the second-best seller. Memory expands buyers' choice sets and, in doing so, reduces the gap between the best and the second-best seller in the buyers' choice sets, which lowers seller profits and increases buyers' surplus. Word of mouth does not expand the buyers' choice sets. Word of mouth tilts the content of part of the buyers' choice sets towards better sellers. This increases concentration but it does not increase competition for *all* sellers in the way that an increase in the size of buyers' choice sets does. The effect is more complex: competition increases for the best seller but decreases for the weakest sellers.

Lastly, we want to compare the equilibrium of the model with word of mouth ($\beta > 0$) with the equilibrium of a model without word of mouth ($\beta = 0$), and the equilibrium of the model with more word of mouth (higher β) or less word of mouth (lower β). The results in Proposition 12 follow from (3.16) plus (3.20) and (3.15).

Proposition 12: (Word of mouth). *Take any β_1 and β_2, with $0 \leq \beta_1 < \beta_2 \leq 1$.*

(i) *In period 1, equilibrium surplus, prices and transactions are the same for β_1 and β_2, in the sense that $F_1(s|\beta_1) = F_1(s|\beta_2)$, $G_1(p|\beta_1) = G_1(p|\beta_2)$ and $H_1(y|\beta_1) = H_1(y|\beta_2)$.*

(ii) *In period $t \geq 2$, equilibrium surplus is higher, prices are lower, and transactions are better with more word of mouth (i.e. higher β), in the sense that $F_t(s|\beta_1) > F_t(s|\beta_2)$, $G_t(p|\beta_1) < G_t(p|\beta_2)$ and $H_t(y|\beta_1) > H_t(y|\beta_2)$.*

(iii) *Let $\beta_1 = 0$. Without word of mouth, equilibrium surplus, prices and transactions are constant over time, in the sense that $F_t(s|\beta_1) = F_1(s|\beta_1)$, $G_t(p|\beta_1) = G_1(p|\beta_1)$ and $H_t(y|\beta_1) = H_1(y|\beta_1)$.*

We have seen that prices may be either increasing, decreasing, or non-monotonic over time for different sellers with word of mouth, depending on the seller's quality and popularity. We can also consider the dynamics of the *average* price $\tilde{p}_t = E_{H_t}(p_t(y))$ and the *average* surplus $\tilde{s}_t = E_{H_t}(s_t(y))$ across transactions. We define the *consumer surplus* at time t as the average surplus received by a consumer:

$$cs_t = \underbrace{(1 - e^{-\lambda_t})}_{\textit{prob. transaction}} \underbrace{\tilde{s}_t}_{\textit{avge surplus}},$$

i.e. the probability a buyer makes a transaction in period t multiplied by the average surplus per transaction in that period.

We can decompose the consumer surplus into three components: the probability $1 - e^{-\lambda_t}$ of making a transaction, the average transaction quality $\tilde{y}_t = E_{H_t}(y)$, and the surplus share of consumers $\tilde{s}_t/\tilde{y}_t$, i.e. the average surplus divided by the average quality across transactions:

$$cs_t = \underbrace{(1 - e^{-\lambda_t})}_{\textit{prob. transaction}} \underbrace{\tilde{y}_t}_{\textit{avge quality}} \underbrace{\frac{\tilde{s}_t}{\tilde{y}_t}}_{\textit{surplus share}}.$$

With memory, the consumer surplus is always increasing over time. Surprisingly, however, the consumer surplus may be either increasing or decreasing over time with word of mouth.

Figure 3 illustrates an example of the dynamics of the consumer surplus and the consumer surplus share over time. In this example, the consumer surplus is increasing over time for both memory and word and mouth. The dynamics of the consumer surplus share look similar for both. However, if we look more closely, we can see that the dynamics of the consumer surplus share $\tilde{s}_t/\tilde{y}_t$ are *non-monotonic* with word of mouth, but not with memory. Despite the average quality of transactions increasing over time, which increases

the overall consumer surplus over time, the surplus share of consumers is increasing at first and then decreasing over time with word of mouth. This means that not only can competition vary non-monotonically over time for specific sellers, but aggregate competition – as measured by the surplus share of consumers – can be non-monotonic over time.

3.4 Steady state

With word of mouth, the average number of contacts $\lambda = \mu + \rho$ does not vary over time, so the limiting behavior of the product market is determined by the limiting behavior of the sampling distribution $\Gamma_t(y)$, which is determined in turn by the limiting behavior of the transaction distribution $H_t(y)$. To discuss this limiting behavior, we need to first establish that there is convergence of $H_t(y)$ and therefore of $\Gamma_t(y)$ to an ergodic distribution.

Proposition 13 says that if the average proportion of contacts β obtained through word of mouth is sufficiently low, the transaction distribution converges to an atomless distribution $H^*(y)$.

Proposition 13: (Ergodic transaction distribution with word of mouth) *If* $\beta < \frac{1-e^{-\lambda}}{\lambda}$, *we have the following:*

(i) *The transaction distribution* $H_t(y)$ *converges to an ergodic distribution* $H^*(y)$ *that is given by*

$$H^*(y) = \frac{e^{-\lambda(1-\Gamma^*(y))} - e^{-\lambda}}{1 - e^{-\lambda}} \tag{3.40}$$

on $[y_\ell, y_h]$ *and the distribution* $H^*(y)$ *is atomless.*

(ii) *The sampling distribution* $\Gamma_t(y)$ *converges to an ergodic distribution* $\Gamma^*(y)$ *given by*

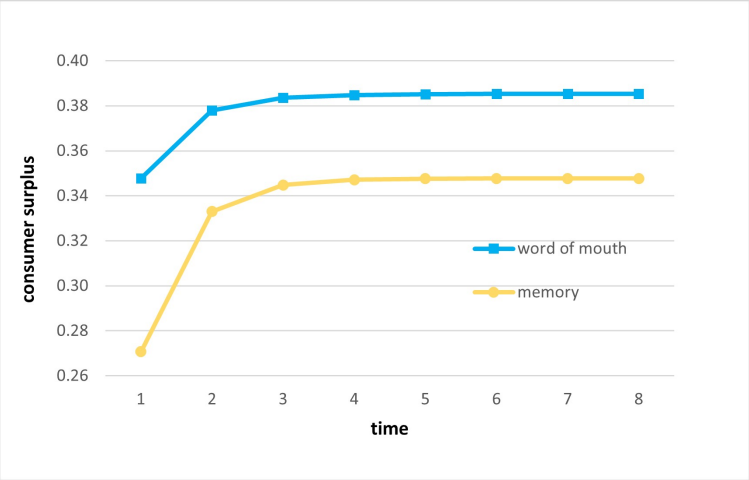
$$\Gamma^*(y) = \beta H^*(y) + (1-\beta)\Phi(y) \tag{3.42}$$

where $H^*(y) < \Gamma^*(y) < \Phi(y)$ *for any* $y \in (y_\ell, y_h)$.

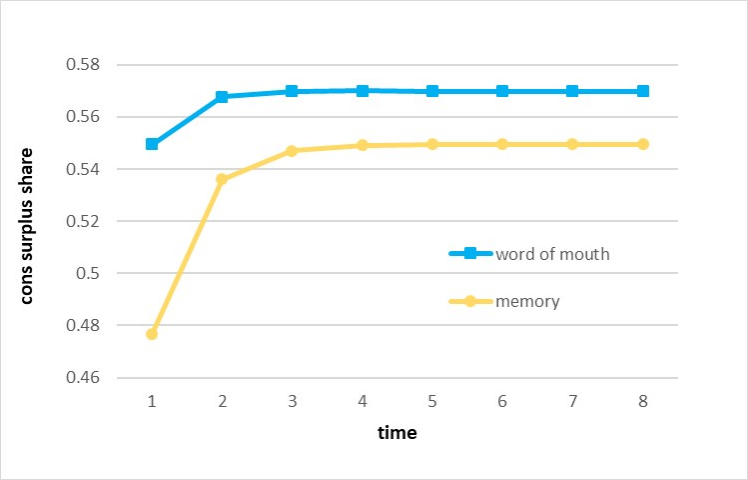
If $\beta < \frac{1-e^{-\lambda}}{\lambda}$, then for any $y < y_h$, the surplus function $s_t(y)$ is eventually strictly decreasing over time, i.e. there exists a time τ such that $s_{t+1}(y) < s_t(y)$ for all $t > \tau$. To see this, the surplus function $s_t(y)$ is decreasing at time t, i.e. $s_{t+1}(y) < s_t(y)$, if $y < y_t^c$ by Proposition 8. Also, Proposition 13 says that $y^c_{\max} = \lim_{t\to\infty} y_t^c = y_h$, which implies that, eventually, for any $y < y_h$, there will be some time τ at which $y < y_\tau^c$. Since y_t^c is increasing over time, this implies that $y < y_t^c$ and for all $t > \tau$ and therefore $s_t(y)$ is strictly decreasing for all $t > \tau$. This means there is convergence to a steady state surplus function $s^*(y)$ because $s_t(y)$ is bounded, i.e. $s_t(y) \in [0, y]$ for all $y \in [y_l, y_h]$, and $s_t(y)$ is either eventually strictly decreasing over time (if $y < y_h$) or strictly increasing over time (if $y = y_h$). Therefore, $s_t(y) \to s^*(y)$ for all $y \in [y_l, y_h]$.

The steady state surplus function $s^*(y)$ is given by the solution to the differential

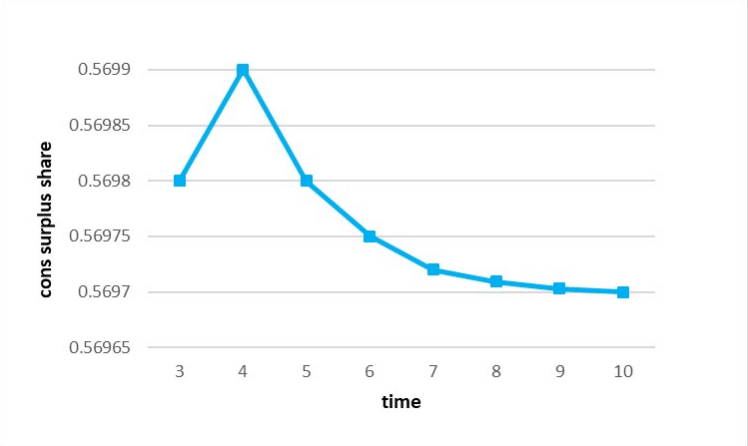
Figure 3: Dynamics of consumer surplus with $\rho = 0.5$, $\mu = 2$, $\delta = 0.8$, uniform $\Phi(y)$



(a) Dynamics of consumer surplus over time



(b) Dynamics of consumer surplus share over time



(c) Non-monotonic consumer surplus share with word of mouth

equation

$$s^{*\prime}(y) = \Phi'(y)\left(\mu + \rho\kappa^*(y)\right)\left(y - s^*(y)\right), \tag{3.43}$$

where the popularity of any seller with quality y is

$$\kappa^*(y) = \frac{H^{*\prime}(y)}{\Phi'(y)}. \tag{3.44}$$

By similar reasoning to Theorem 2, we obtain the following.

Proposition 14: (Existence and uniqueness of stationary equilibrium with word of mouth) *If* $\beta < \frac{1-e^{-\lambda}}{\lambda}$*, an equilibrium exists and is unique. The steady state surplus function is*

$$s^*(y) = \frac{\int_{y_\ell}^{y} \lambda e^{-\lambda(1-\Gamma^*(x))} x \Gamma^{*\prime}(x) dx}{e^{-\lambda(1-\Gamma^*(y))}} \tag{3.49}$$

for all $y \in [y_\ell, y_h]$ *and* $s^*(y_h) = (1 - e^{-\lambda})\tilde{y}^*$ *where* $\tilde{y}^*$ *is the average transaction quality.*

We now compare sales concentration in the stationary limit of the model with memory and the model with word of mouth where $\beta < \frac{1-e^{-\lambda}}{\lambda}$. To prevent confusion, let λ_m^* denote the steady state average number of contacts in the model with memory. Suppose that the average number of new sellers contacted through search, μ, is the same for both models. We have $\lambda_m^* = \mu/\delta$ and $\lambda = \mu + \rho$. Now suppose that the average number of contacts per buyer is equal in the steady state, $\lambda_m^* = \lambda$.

The ratio of sales concentration with word of mouth, $Q^*(x)$, and memory, $Q_m^*(x)$, is

$$\frac{Q^*(x)}{Q_m^*(x)} = \frac{1 - e^{-\mu x} e^{-\rho(1-H^*(\Phi^{-1}(1-x)))}}{1 - e^{-(\mu+\rho)x}}. \tag{3.51}$$

We have $Q^*(x) > Q_m^*(x)$ for all $x \in (0,1)$ if and only if

$$\mu x + \rho(1 - H^*(\Phi^{-1}(1-x))) > (\mu + \rho)x. \tag{3.52}$$

Rearranging, this holds if and only if

$$H^*(\Phi^{-1}(1-x)) < 1 - x, \tag{3.53}$$

In general, $H^*(y) < \Gamma^*(y) < \Phi(y)$ for any $\rho > 0$ and $y \in (y_\ell, y_h)$, so in particular

$$H^*(\Phi^{-1}(1-x)) < \Phi(\Phi^{-1}(1-x)) = 1 - x. \tag{3.54}$$

Thus sales concentration is higher with word of mouth than with memory if the average number of contacts is the same. We summarize this observation in Proposition 15.

Proposition 15: (Concentration with memory vs word of mouth) *Suppose the steady state average number of contacts* λ *is the same with memory and with word of mouth. If* $\beta < \frac{1-e^{-\lambda}}{\lambda}$*, sales concentration in the steady state* $Q^*(x)$ *is strictly higher with word of mouth than with memory for any* $x \in (0,1)$*.*

In contrast to the model with memory, where competition is increasing over time for all sellers, competition does not increase over time for all sellers with word of mouth. However, competition always increases over time for the top seller, which has quality y_h. We might therefore wonder whether the economy converges to a competitive outcome at least for the top seller. If $\beta < \frac{1-e^{-\lambda}}{\lambda}$, the surplus offered by the top seller is $s^*(y_h) = (1 - e^{-\lambda})\tilde{y}^*$ where $\tilde{y}^*$ is the average transaction quality, so we have $s^*(y_h) < y_h$ and $p^*(y_h) > 0$ for any finite λ. Only in the frictionless limit as $\lambda^* \to \infty$ is the equilibrium perfectly competitive. As $\lambda^* \to \infty$, the distribution of transactions $H^*(y)$ is degenerate at y_h and we have $s^*(y_h) \to \tilde{y}^* = y_h$ and $p^*(y_h) \to 0$.

4 Conclusion

This paper develops a tractable framework for dynamic learning in the static search-theoretic model of imperfect competition developed in Butters (1977), Varian (1980) and Burdett and Judd (1983). In this model of imperfect competition, buyers can only purchase from a subset of sellers in each period, called their "choice set", due to the presence of search frictions. With learning, either the average *size* of buyers' choice sets can increase over time or the *composition* of buyers' choice sets can change over time.

We consider both private learning (via memory) and social learning (via word of mouth). Both forms of learning are highly tractable. We establish the existence and uniqueness of equilibrium and characterize the dynamics of the product market for both models. Overall, we find that the dynamics of market concentration are similar for both memory and word of mouth, but the dynamics of competition and prices are different.

Memory increases the quantity of information available to buyers by increasing the *size* of their choice sets over time. As a result, memory leads to both higher market concentration over time and stronger competition over time for all sellers.

Word of mouth, on the other hand, increases the quality of information available to buyers by changing the *composition* of the subset of sellers in their choice sets. Over time, buyers' choice sets are more likely to include higher-quality sellers. As a result, word of mouth leads to higher concentration over time. However, unlike memory, word of mouth does not increase competition for *all* sellers. The dynamics of competition and prices are heterogeneous across sellers and depend on a seller's popularity. At any given point in time, we find that competition is increasing and prices are decreasing over time for higher-quality sellers, but the opposite is true for lower-quality sellers. Competition and prices may also vary non-monotonically over time for sellers of intermediate quality.

Given the tractability of our framework for dynamic learning, we believe there are many potential applications of the model – both for product markets and the labor market.

References

[1] Albrecht, J., G. Menzio, and S. Vroman. 2023. "Vertical Product Differentiation in Frictional Markets." *Journal of Political Economy: Macro*, 1: 586-632.

[2] Banerjee, A. and D. Fudenberg. 2004. "Word-of-mouth Learning. " *Games and Economic Behavior*, 46(1): 1-22.

[3] Burdett, K., and K. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica*, 51: 955-970.

[4] Burdett, K., A. Trejos and R. Wright. 2017. "A New Suggestion for Simplifying the Theory of Money." *Journal of Economic Theory*, 172: 423-450.

[5] Burdett, K., and G. Menzio. 2018. "The (Q,S,s) Pricing Rule." *Review of Economic Studies*, 85: 892-928.

[6] Burdett, K., and D. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39: 955-970.

[7] Butters, G. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *Review of Economic Studies*, 44: 465-491.

[8] Cai, X., P. Gautier, and R. Wolthoff. 2023. "Spatial Search." Manuscript, University of Toronto.

[9] Campbell, A. 2013. "Word-of-mouth Communication and Percolation in Social Networks." *American Economic Review*, 103(6): 2466-2498.

[10] Campbell, A. 2015. "Word of Mouth Model of Sales." *Economics Letters*, 133: 45-50.

[11] Campbell, A., C.M. Leister and Y Zenou. 2020. "Word-of-mouth Communication and Search." *The RAND Journal of Economics*, 51(3): 676-712.

[12] Campbell, A., P. Ushchev and Y. Zenou. 2023. "The Network Origins of Entry." *Journal of Political Economy*, forthcoming.

[13] Ellison, G., and D. Fudenberg. 1995. "Word-of-mouth Communication and Social Learning. " *The Quarterly Journal of Economics*, 110(1): 93-125.

[14] Galeotti, A. 2010. "Talking, Searching, and Pricing." *International Economic Review*, 51(4): 1159-1174.

[15] Galenianos, M., R. Pacula, and N. Persico. 2012. "A Search-Theoretic Model of the Retail Market for Illicit Drugs." *Review of Economic Studies*, 79: 1239-1269.

[16] Galenianos, M. 2013. "Learning about Match Quality and the Use of Referrals." *Review of Economic Dynamics*, 16(4): 668-690.

[17] Galenianos, M. 2014. "Hiring through Referrals." *Journal of Economic Theory*, 152: 304-323.

[18] Galenianos, M. 2021. "Referral Networks and Inequality." *The Economic Journal*, 131(633): 271-301.

[19] Head, A. and A. Kumar. 2005. "Price dispersion, Inflation, and Welfare." *International Economic Review*, 46(2): 533-572.

[20] Head, A., L. Liu, G. Menzio, and R. Wright. 2012. "Sticky Prices: A New Monetarist Approach." *Journal of the European Economic Association*, 10: 939-973.

[21] Kaplan, G., and G. Menzio. 2015. "The Morphology of Price Dispersion." *International Economic Review*, 56: 892-928.

[22] Kaplan, G., and G. Menzio. 2016. "Shopping Externalities and Self-Fulfilling Unemployment Fluctuations." *Journal of Political Economy*, 124: 771-825.

[23] Kaplan, G., G. Menzio, L. Rudanko, and N. Trachter. 2019. "Relative Price Dispersion: Evidence and Theory." *American Economic Journal: Micro*, 11: 68-124.

[24] Menzio, G. 2023. "Optimal Product Design: Implications for Competition and Growth of Declining Search Frictions." *Econometrica*, 91: 605-639.

[25] Menzio, G. 2023. "Search Theory of Imperfect Competition with Decreasing Returns to Scale." NBER Working Paper 31174.

[26] Menzio, G. 2024. "Markups: A Search-Theoretic Perspective. " NBER Working Paper 32888.

[27] Nord, L. 2022. "Shopping, Demand Composition, and Equilibrium Prices." Manuscript, European University Institute.

[28] Pytka, K. 2018. "Shopping Effort in Self-Insurance Economies." Manuscript, University of Mannheim.

[29] Shi, S. 2023. "Sequentially Mixed Search and Equilibrium Price Dispersion." *Journal of Economic Theory*, 213: 105735

[30] Sorensen, A. 2000. "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs." *Journal of Political Economy*, 108: 833-862.

[31] Varian, H. 1980. "A Model of Sales." *American Economic Review*, 61: 37-70.

Appendix: Proofs

Proofs for Section 2

Proof of Lemma 1: Suppose that the buyer enters the period with n contacts, where n is Poisson with some arbitrary coefficient $\gamma > 0$. Moreover, suppose that the buyer maintains a contact with probability $1-\delta$ and loses a contact with probability δ, with $\delta \in [0,1]$. Let $\hat{n}$ denote the number of contacts retained by the buyer. The probability that $\hat{n}$ equals $k = 0, 1, 2, ...n$ is given by

$$\begin{aligned}
\Pr(\hat{n}=k) &= \sum_{n=k}^{\infty} \frac{e^{-\gamma}\gamma^n}{n!} \frac{n!}{k!(n-k)!}(1-\delta)^k \delta^{n-k} \\
&= \frac{e^{-\gamma}e^{\gamma\delta}\gamma^k(1-\delta)^k}{k!}\left[\sum_{n=k}^{\infty} \frac{\gamma^{n-k}}{1}\frac{e^{-\gamma\delta}}{(n-k)!}\delta^{n-k}\right] \\
&= \frac{e^{-\gamma(1-\delta)}(\gamma(1-\delta))^k}{k!}.
\end{aligned} \tag{A.1}$$

The expression in the first line of (A.1) is easy to understand. The probability that $\hat{n} = k$ is equal to the sum of the probability that the buyer enters the market with n contacts and retains k of them, for $n = k, k+1, k+2,....$ The first term on the right-hand side of (A.1) is the probability that the buyer enters the market with n contacts, given that n is distributed as a Poisson with coefficient γ. The second term is the probability that the buyer maintains exactly k of its n contacts, given that the buyer retains a contact with probability δ. The second line of (A.1) is obtained by collecting terms. The third line of (A.1) follows from the fact that the summation in the second line is equal to 1. The third line reveals that $\hat{n}$ is distributed like a Poisson with coefficient $\gamma(1-\delta)$.

Next, suppose that the buyer retains a number $\hat{n}$ of old contacts, where $\hat{n}$ is Poisson with some arbitrary coefficient $\hat{\gamma} > 0$. Moreover, suppose that the buyer makes m new contacts, where m is Poisson with coefficient $\mu > 0$. Let n_+ denote the total number of contacts of the buyer. The probability that n_+ equals k is given by

$$\begin{aligned}
\Pr(n_+=k) &= \sum_{\hat{n}=0}^{k} \frac{e^{-\hat{\gamma}}\hat{\gamma}^{\hat{n}}}{\hat{n}!}\frac{e^{-\mu}\mu^{k-\hat{n}}}{(k-\hat{n})!} \\
&= \frac{e^{-(\hat{\gamma}+\mu)}(\hat{\gamma}+\mu)^k}{k!}\left[\sum_{\hat{n}=0}^{k} \frac{k!}{\hat{n}!(k-\hat{n})!}\left(\frac{\hat{\gamma}}{\hat{\gamma}+\mu}\right)^{\hat{n}}\left(\frac{\mu}{\hat{\gamma}+\mu}\right)^{k-\hat{n}}\right] \\
&= \frac{e^{-(\hat{\gamma}+\mu)}\left(\hat{\gamma}+\mu\right)^k}{k!}.
\end{aligned} \tag{A.2}$$

The expression in the first line of (A.2) is easy to understand. The probability that the buyer has k total contacts is equal to the sum of the probability that the buyer retains $\hat{n}$ contacts and adds $k-\hat{n}$ new contacts, for $\hat{n} = 0, 1, 2....k$. The first term on the right-hand side of (A.2) is the probability that the buyer retains $\hat{n}$ contacts, given that $\hat{n}$ is

distributed as a Poisson with coefficient $\hat{\gamma}$. The second term is the probability that the buyer adds $m = k - \hat{n}$ contacts, given that m is Poisson with coefficient μ. The second line of (A.2) is obtained by collecting terms. The third line of (A.2) follows from the fact that the summation in the second line is equal to 1. The third line reveals that n_+ is distributed like a Poisson with coefficient $\hat{\gamma} + \mu$.

By assumption, n_1 is Poisson with coefficient $\lambda_1 = \mu$. From the above observations, it follows that $\hat{n}_2$ is Poisson with coefficient $\lambda_1(1-\delta)$ and, in turn, n_2 is Poisson with coefficient $\lambda_2 = \lambda_1(1-\delta) + \mu$. Similarly, it follows that $\hat{n}_3$ is Poisson with coefficient $\lambda_2(1-\delta)$ and, in turn, n_3 is Poisson with coefficient $\lambda_3 = \lambda_2(1-\delta) + \mu$. In general, $\hat{n}_t$ is Poisson with coefficient $\lambda_{t-1}(1-\delta)$, and n_t is Poisson with coefficient $\lambda_t = \lambda_{t-1}(1-\delta) + \mu$.

We prove that n_t is distributed as a Poisson with coefficient $\lambda_t = \mu \sum_{i=1}^{t} (1-\delta)^{t-i}$ by induction on t. First, n_1 is distributed as a Poisson with coefficient $\lambda_1 = \mu$. Now suppose that n_t is distributed as a Poisson with coefficient $\lambda_t = \mu \sum_{i=1}^{t} (1-\delta)^{t-i}$. Consider period $t+1$. We know that n_{t+1} is distributed as a Poisson with coefficient $\lambda_{t+1} = \lambda_t(1-\delta) + \mu$. Therefore, since we assume that $\lambda_t = \mu \sum_{i=1}^{t} (1-\delta)^{t-i}$, we have $\lambda_{t+1} = \mu \sum_{i=1}^{t} (1-\delta)^{t-i}(1-\delta) + \mu$ or $\lambda_{t+1} = \mu \left(\sum_{i=1}^{t} (1-\delta)^{t+1-i} + 1 \right)$. Simplifying delivers $\lambda_{t+1} = \mu \sum_{i=1}^{t+1} (1-\delta)^{t+1-i}$. Therefore, n_t is distributed as a Poisson with coefficient $\lambda_t = \mu \sum_{i=1}^{t} (1-\delta)^{t-i}$ for all $t \geq 1$. Rearranging yields $\lambda_t = \mu \sum_{i=0}^{t-1} (1-\delta)^{i}$, which is equal to $\mu \frac{1-(1-\delta)^t}{1-(1-\delta)} = \frac{\mu}{\delta}[1-(1-\delta)^t]$. ■

Proof of Lemma 2: If a seller with quality $y \in [y_\ell, y_h]$ offers a surplus s equal to 0, its profit is

$$\begin{aligned} V_t(y,0) &= \left[\sum_{k=0}^{\infty} b_{k,t} \pi_{k,t}(0)\right] y \\ &\geq b_{0,t} \pi_{0,t}(0) y. \end{aligned} \tag{A.3}$$

It is clear from (A.3) that $b_{0,k} > 0$. That is, the seller meets a strictly positive measure of buyers that are captive. It is clear from (A.3) that $\pi_{0,t}(0) = 1$. That is, a captive buyer purchases the good from the seller with probability 1. Lastly, since $y \in [y_\ell, y_h]$ and $y_\ell > 0$, the seller makes a strictly positive profit per sale by offering a surplus of 0. These observations imply that the second line in (A.3) is strictly positive and so is $V_t(y,0)$. Since offering a surplus of 0 is feasible but not necessarily optimal, it follows that the seller's maximized profit is strictly positive. ■

Proof of Lemma 3: On the way to a contradiction, suppose that the distribution $F_t(s)$ has a mass point at some surplus s_0. Let y_0 denote the quality of one of the sellers offering s_0. From Lemma 2, it follows that $y_0 > s_0$. By offering the surplus s_0, the seller enjoys a profit of

$$V_t(y_0, s_0) = \left[\sum_{k=0}^{\infty} b_{k,t} \pi_{k,t}(s_0)\right] (y_0 - s_0). \tag{A.4}$$

By offering a surplus of $s_0 + \epsilon$, with $\epsilon > 0$, the seller enjoys a profit of

$$\begin{aligned} & V_t(y_0, s_0 + \epsilon) \\ = \; & \left[\sum_{k=0}^{\infty} b_{k,t} \pi_{k,t}(s_0 + \epsilon)\right](y_0 - s_0 - \epsilon) \\ \geq \; & \left[\sum_{k=0}^{\infty} b_{k,t} \pi_{k,t}(s_0)\right](y_0 - s_0 - \epsilon) \\ + \; & \left[\sum_{k=0}^{\infty} b_{k,t} \left(\sum_{j=1}^{k} \frac{j!(k-j)!}{k!} \zeta_t(s)^j F_t(s-)^{k-j} \left(1 - \frac{1}{j+1}\right)\right)\right](y_0 - s_0 - \epsilon), \end{aligned} \tag{A.5}$$

where the inequality in the third line follows from the fact that buyers that have j other contacts offering s_0 and $k - j$ other contacts offering strictly less than s_0 purchase the good from the seller with probability less than 1 if the seller offers s_0 and with probability 1 if the seller offers $s_0 + \epsilon$. Since $\zeta_t(s) > 0$, the difference in trading probability in the fourth line of (A.5) is strictly positive. Also, the difference in the trading probability in the fourth line is independent of ϵ. Since $s_0 < y_0$, there is some ϵ small enough such that $V_t(y_0, s_0 + \epsilon)$ is strictly greater than $V_t(y_0, s_0)$. Since a seller must maximize its profit and $V_t(y_0, s_0 + \epsilon) > V_t(y_0, s_0)$, a seller with quality y_0 cannot offer the surplus s_0. Therefore, $F_t(s)$ cannot be an equilibrium distribution. ■

Proof of Lemma 4: We first establish that the support of F_t is an interval $[s_{\ell,t}, s_{h,t}]$. On the way to a contradiction, suppose that the support of F_t has a gap between s_0 and s_1, with $s_0 < s_1$, and s_0 and s_1 both on the support of F_t. Let y_1 denote the quality of a seller that offers the surplus s_1. The profit of this seller is given by

$$V_t(y_1, s_1) = b\lambda_t e^{-\lambda_t(1 - F_t(s_1))}(y_1 - s_1). \tag{A.6}$$

If the seller offers the surplus s_0, the seller's profit is given by

$$\begin{aligned} V_t(y_1, s_0) &= b\lambda_t e^{-\lambda_t(1 - F_t(s_0))}(y_1 - s_0) \\ &= b\lambda_t e^{-\lambda_t(1 - F_t(s_1))}(y_1 - s_0) \\ &> b\lambda_t e^{-\lambda_t(1 - F_t(s_1))}(y_1 - s_1) = V_t(y_1, s_1), \end{aligned} \tag{A.7}$$

where the second line follows from the fact that $F(s_1) = F(s_0)$, and the third line follows from the fact that $s_0 < s_1$. Since a seller must maximize its profit and $V_t(y_1, s_0) > V_t(y_1, s_1)$, a seller with quality y_1 cannot offer the surplus s_1. Therefore, $F_t(s)$ cannot be an equilibrium distribution.

Next, we establish that $s_{\ell,t} = 0$. On the way to a contradiction, suppose $s_{\ell,t} < 0$. Let y_0 denote the quality of a seller that offers the surplus $s_{\ell,t}$. The seller attains a profit $V_t(y_0, s_{\ell,t})$ equal to 0, since no buyer purchases the good for a strictly negative surplus. If, however, the seller were to offer a surplus of 0, it would attain a profit $V_t(y_0, 0)$, which is strictly positive by Lemma 1. Therefore, $s_{\ell,t}$ cannot be strictly negative. Next, suppose $s_{\ell,t} > 0$. Let y_0 denote the quality of a seller that offers the surplus $s_{\ell,t}$. The seller attains a profit $V_t(y_0, s_{\ell,t})$ equal to $b\lambda_t \exp(-\lambda_t)(y_0 - s_{\ell,t})$. If, however, the seller were to offer a

surplus of 0, it would attain a profit $V_t(y_0, 0)$ equal to $b\lambda_t \exp(-\lambda_t)y_0$, which is strictly greater than $V_t(y_0, s_{\ell,t})$. Therefore, $s_{\ell,t}$ cannot be strictly positive. Combining the above observations yields $s_{\ell,t} = 0$. ■

Proof of Lemma 5: The proof has two parts. The first part establishes that surplus offered by a seller is strictly increasing in the seller's quality. The second part establishes that the mapping between the seller's quality and the surplus is a function, in the sense that every seller with the same quality offers the same surplus.

(i) Consider a seller with quality y_0 and a seller with quality y_1, with $y_0 < y_1$. Denote as s_0 the surplus offered by the seller with quality y_0, and denote as s_1 the surplus offered by the seller with quality y_1. Since the seller with quality y_0 enjoys a higher profit offering s_0 rather than s_1, we have

$$b\lambda_t e^{-\lambda(1-F_t(s_0))}(y_0 - s_0) \geq b\lambda_t e^{-\lambda(1-F_t(s_1))}(y_0 - s_1). \tag{A.8}$$

Since the seller with quality y_1 enjoys a higher profit offering s_1 rather than s_0, we have

$$b\lambda_t e^{-\lambda_t(1-F_t(s_1))}(y_1 - s_1) \geq b\lambda_t e^{-\lambda_t(1-F_t(s_0))}(y_1 - s_0). \tag{A.9}$$

Combining (A.8) and (A.9) yields

$$\left[e^{-\lambda_t(1-F_t(s_1))} - e^{-\lambda_t(1-F_t(s_0))}\right](y_1 - y_0) \geq 0. \tag{A.10}$$

Since $y_1 > y_0$ and $F_t(s)$ is strictly increasing in s, (A.10) implies that s_1 is greater or equal to s_0. Now, suppose that $s_0 = s_1 = s$. Then, every seller with quality $y \in (y_0, y_1)$ must offer the surplus s and, hence, there is a mass point in the surplus distribution F_t, which contradicts Lemma 3. Hence, it must be the case that s_1 is strictly greater than s_0.

(ii) On the way to a contradiction, suppose that there is a seller with quality y_0 that offers the surplus s_0 and another seller with quality y_0 that offers the surplus s_1, with $s_0 < s_1$. Every seller with quality $y > y_0$ must offer a surplus strictly greater than s_1. Every seller with quality $y < y_0$ must offer a surplus strictly smaller than s_0. Therefore, $F_t(s_0)$ must be equal to $F_t(s_1)$ or, in other words, there must be a gap on the support of F_t between s_0 and s_1, which contradicts Lemma 4. ■

Proof of Theorem 1: We have already established the existence and uniqueness of equilibrium. Rewriting (2.11), the function $s_t(y)$ solves the following differential equation:

$$s_t'(y) + \lambda_t \Phi'(y) s_t(y) = \lambda_t y \Phi'(y), \tag{A.11}$$

where $s_t(y_\ell) = 0$. For any given y, there exists a unique solution $s_t(y)$ which satisfies this differential equation and the general form of the solution is

$$s_t(y) = \lambda_t e^{-\lambda_t \Phi(y)} \int_{y_\ell}^{y} e^{\lambda_t \Phi(x)} x \Phi'(x) dx + c e^{-\lambda_t \Phi(y)}, \tag{A.12}$$

where c is a constant. Given that $s_t(y_\ell) = 0$, we have $c e^{-\lambda_t \Phi(y_\ell)} = c = 0$. Therefore,

$$s_t(y) = \int_{y_\ell}^{y} \lambda_t e^{-\lambda_t (\Phi(y) - \Phi(x))} x \Phi'(x) dx. \tag{A.13}$$

Rearranging, (A.13) is equivalent to (2.12).

To prove that $s_t(y_h) = (1 - e^{-\lambda_t}) \tilde{y}_t$, note that the distribution of qualities of transactions $H_t(y)$ is given by (2.6) and the average quality of a transaction, $\tilde{y}_t = E_{H_t}(y)$, is given by (2.7). To calculate $s_t(y_h)$, expression (A.13) for $s_t(y)$ implies

$$s_t(y_h) = \int_{y_\ell}^{y_h} \lambda_t e^{-\lambda_t (1 - \Phi(x))} x \Phi'(x) dx. \tag{A.14}$$

Therefore, we have $s_t(y_h) = (1 - e^{-\lambda_t}) \tilde{y}_t$. ■

Proof of Proposition 2: For $\gamma > 0$, let $h(\gamma)$ be defined as

$$h(\gamma) = \frac{e^{-\gamma(1-x)} - e^{-\gamma}}{1 - e^{-\gamma}}. \tag{A.15}$$

The derivative of $h(\gamma)$ with respect to γ is given by

$$h'(\gamma) = \frac{e^{-\gamma} \left[1 - x e^{-\gamma(1-x)} - (1-x) e^{\gamma x} \right]}{(1 - e^{-\gamma})^2}. \tag{A.16}$$

The term in square brackets at the numerator of (A.16) determines the sign of $h'(\gamma)$. The term takes the value 0 for $x = 0$, it takes the value 0 for $x = 1$, and it is strictly convex with respect to x for all $x \in (0, 1)$. Therefore, the term in square brackets is strictly negative for all $x \in (0, 1)$ and, in turn, $h'(\gamma)$ is strictly negative for all $x \in (0, 1)$. Since $H_t(y) = h(\lambda_t)$ and $H_{t+1}(y) = h(\lambda_{t+1})$ for $x = \Phi(y)$ and $\lambda_t < \lambda_{t+1}$, it follows that $H_{t+1}(y) < H_t(y)$ for all $y \in (y_\ell, y_h)$. ■

Proofs for Section 3

Proof of Lemma 6: A buyer purchases the good from the contacted seller that carries the variety of the good with highest quality. The fraction of buyers that purchase the good from a seller with quality non-greater than y is

$$\begin{aligned} & H_t(y) \\ = \; & \left\{ b e^{-\mu} \textstyle\sum_{j=1}^{\infty} \frac{e^{-\rho} \rho^j}{j!} \hat{\Phi}_t(y)^j \right\} \Big/ \left\{ b \left[1 - e^{-(\mu+\rho)} \right] \right\} \\ + \; & \left\{ \textstyle\sum_{k=1}^{\infty} b \frac{e^{-\mu} \mu^k}{k!} \left[\sum_{j=0}^{\infty} \frac{e^{-\rho} \rho^j}{j!} \Phi(y)^k \hat{\Phi}_t(y)^j \right] \right\} \Big/ \left\{ b \left[1 - e^{-(\mu+\rho)} \right] \right\}. \end{aligned} \tag{A.25}$$

Let us explain (A.25). The first term at the numerator is the measure of buyers that do not reach any seller through search and reach j sellers through word of mouth, with $j = 1, 2, 3 \ldots$. Each one of these buyers purchases the good from a seller with quality non-greater than y with probability $\hat{\Phi}_t(y)^j$. The second term at the numerator is the measure of buyers that reach k sellers through search, with $k = 1, 2, \ldots$, and reach j sellers through word of mouth, with $j = 0, 1, 2, \ldots$ Each one of these buyers purchases the good from a seller with quality non-greater than y with probability $\Phi(y)^k \hat{\Phi}_t(y)^j$. Overall, the numerator is the measure of buyers that purchase the good from a seller with quality non-greater than y. The denominator is the measure of buyers who purchase the good, which is equal to the measure of buyers multiplied by the probability that a buyer contacts at least one seller either through search or word of mouth.

We can rewrite the first term at the numerator of (A.25) as

$$\begin{aligned} & be^{-\mu} \textstyle\sum_{j=1}^{\infty} \dfrac{e^{-\rho}\rho^j}{j!} \hat{\Phi}_t(y)^j \\ = \; & be^{-\mu} \left[e^{-\rho(1-\hat{\Phi}_t(y))} \textstyle\sum_{j=0}^{\infty} \dfrac{e^{-\rho\hat{\Phi}_t(y)}\rho^j}{j!} \hat{\Phi}_t(y)^j - e^{-\rho} \right] \\ = \; & be^{-\mu} \left[e^{-\rho(1-\hat{\Phi}_t(y))} - e^{-\rho} \right]. \end{aligned} \tag{A.26}$$

The first equality in (A.26) is obtained by collecting terms. The second equality in (A.26) is obtained by recognizing that the summation over j is equal to 1.

We can rewrite the second term at the numerator of (A.25) as

$$\begin{aligned} & \textstyle\sum_{k=1}^{\infty} b \dfrac{e^{-\mu}\mu^k}{k!} \left[\textstyle\sum_{j=0}^{\infty} \dfrac{e^{-\rho}\rho^j}{j!} \Phi(y)^k \hat{\Phi}_t(y)^j \right] \\ = \; & \textstyle\sum_{k=1}^{\infty} b \dfrac{e^{-\mu}\mu^k \Phi(y)^k}{k!} e^{-\rho(1-\hat{\Phi}_t(y))} \left[\textstyle\sum_{j=0}^{\infty} \dfrac{e^{-\rho\hat{\Phi}_t(y)}\rho^j \hat{\Phi}_t(y)^j}{j!} \right] \\ = \; & be^{-\rho(1-\hat{\Phi}_t(y))} \left[e^{-\mu(1-\Phi(y))} \textstyle\sum_{k=0}^{\infty} \dfrac{e^{-\mu\Phi(y)}\mu^k \Phi(y)^k}{k!} - e^{-\mu} \right] \\ = \; & be^{-\rho(1-\hat{\Phi}_t(y))} \left[e^{-\mu(1-\Phi(y))} - e^{-\mu} \right]. \end{aligned} \tag{A.27}$$

The first equality in (A.27) is obtained by collecting terms. The second equality in (A.27) is obtained by recognizing that the summation over j is equal to 1. The last equality in (A.27) is obtained by recognizing that the summation over k is equal to 1.

Using the expressions in (A.26) and (A.27) to replace the first and the second term at the numerator of (A.25), and then setting $\hat{\Phi}_t(y) = H_{t-1}(y)$, yields expression (3.15) for $H_t(y)$ in Lemma 10.

To prove that $H_t(y)$ is an atomless distribution, it remains to show that

$$H_t'(y) = \frac{d}{dy}\frac{e^{-\mu(1-\Phi(y))}e^{-\rho(1-H_{t-1}(y))} - e^{-(\mu+\rho)}}{1-e^{-(\mu+\rho)}} > 0, \tag{A.28}$$

where

$$H_t'(y) = \frac{(\mu+\rho\kappa_t(y))\,\Phi'(y)e^{-\mu(1-\Phi(y))}e^{-\rho(1-H_{t-1}(y))}}{1-e^{-(\mu+\rho)}}. \tag{A.29}$$

First, in period 1, we have

$$H_1'(y) = \frac{\lambda\Phi'(y)e^{-\lambda(1-\Phi(y))}}{1-e^{-\lambda}} > 0. \tag{A.30}$$

Next, we have $H_t'(y) > 0$ if and only if $\mu + \rho\kappa_t(y) > 0$, which is true if $H_{t-1}'(y) > 0$ since $\Phi'(y) > 0$. Therefore, by induction, $H_t'(y) > 0$ for all periods $t \geq 1$. ■

Proof of Theorem 2: We have already established existence and uniqueness of equilibrium. Letting $\lambda = \mu + \rho$ and $\beta = \rho/\lambda$, we can rewrite (3.16) as

$$s_t'(y) = \Gamma_t'(y)\lambda(y - s_t(y)) \tag{A.31}$$

and $s_t(y_\ell) = 0$. By similar reasoning to Theorem 1 above, the solution is

$$s_t(y) = \lambda\int_{y_\ell}^{y} e^{-\lambda(\Gamma_t(y)-\Gamma_t(x))}x\Gamma_t'(x)dx. \tag{A.32}$$

To calculate $s_t(y_h)$, we have

$$s_t(y_h) = \lambda\int_{y_\ell}^{y_h} e^{-\lambda(1-\Gamma_t(x))}x\Gamma_t'(x)dx. \tag{A.33}$$

So $s_t(y_h) = (1-e^{-\lambda})\tilde{y}_t$ where

$$\tilde{y}_t = \frac{\int y\lambda\Gamma_t'(x)e^{-\lambda(1-\Gamma_t(x))}dy}{1-e^{-\lambda}}, \tag{A.34}$$

the average transaction quality. ■

Proof of Proposition 6: It follows from (3.22) that $H_{t+1}(y) < H_t(y)$ if and only if

$$\frac{e^{-\lambda(1-\Gamma_{t+1}(y))} - e^{-\lambda}}{1-e^{-\lambda}} < \frac{e^{-\lambda(1-\Gamma_t(y))} - e^{-\lambda}}{1-e^{-\lambda}}, \tag{A.35}$$

which is is true if and only if $e^{-\lambda(1-\Gamma_{t+1}(y))} < e^{-\lambda(1-\Gamma_t(y))}$. This holds if and only if $\Gamma_{t+1}(y) < \Gamma_t(y)$, which is true if and only if $H_t(y) < H_{t-1}(y)$. The result that $H_{t+1}(y) < H_t(y)$ for all t follows by induction starting with $H_2(y) < H_1(y)$. The result that $\Gamma_{t+1}(y) < \Gamma_t(y)$ follows from the fact that $\Gamma_{t+1}(y) - \Gamma_t(y) = \beta(H_t(y) - H_{t-1}(y)) < 0$. ■

Proof of Lemma 7:

(i) Starting with expression (3.17), we have $\kappa_{t+1}(y) > \kappa_t(y)$ if and only if

$$\frac{H_t'(y)}{\Phi'(y)} > \frac{H_{t-1}'(y)}{\Phi'(y)}. \tag{A.36}$$

In particular, $\kappa_2(y) > \kappa_1(y)$ if and only if $\kappa_2(y) > 1$ because $H_0(y) = \Phi(y)$.

In general, since $\Phi'(y) > 0$, we have $\kappa_{t+1}(y) > \kappa_t(y)$ if and only if $H_t'(y) > H_{t-1}'(y)$. Differentiating (3.22) yields

$$H_t'(y) = \frac{\lambda \Gamma_t'(y) e^{-\lambda(1-\Gamma_t(y))}}{1 - e^{-\lambda}}. \tag{A.37}$$

Therefore, we have $\kappa_{t+1}(y) > \kappa_t(y)$ if and only if

$$\frac{\lambda \Gamma_t'(y) e^{-\lambda(1-\Gamma_t(y))}}{1 - e^{-\lambda}} > H_{t-1}'(y). \tag{A.38}$$

Differentiating expression (3.19) for $\Gamma_t(y)$, we obtain

$$\Gamma_t'(y) = \beta H_{t-1}'(y) + (1-\beta)\Phi'(y), \tag{A.39}$$

where $\beta = \rho/\lambda$, so (A.38) is true if and only if

$$\frac{\lambda e^{-\lambda(1-\Gamma_t(y))}}{1 - e^{-\lambda}} > \frac{\lambda}{\frac{\mu}{\kappa_t(y)} + \rho}. \tag{A.40}$$

Now let $f_{L,t}(y)$ denote the left-hand side of (A.40) and $f_{R,t}(y)$ denote the right-hand side of (A.40). Both $f_{L,t}(y)$ and $f_{R,t}(y)$ are strictly increasing in y because $\Gamma_t'(y) > 0$ and $\kappa_t'(y) > 0$. Next, we have $f_{L,t}(y_\ell) < f_{R,t}(y_\ell)$ if and only if $\kappa_{t+1}(y_\ell) < \kappa_t(y_\ell)$, and $f_{L,t}(y_h) > f_{R,t}(y_h)$ if and only if $\kappa_{t+1}(y_h) > \kappa_t(y_h)$. We prove that $\kappa_{t+1}(y_\ell) < \kappa_t(y_\ell)$ and $\kappa_{t+1}(y_h) > \kappa_t(y_h)$ for all $t \geq 1$ by induction.

Using expression (3.25), we have

$$\kappa_t(y) = \frac{(\mu + \rho\kappa_{t-1}(y)) e^{-\mu(1-\Phi(y))} e^{-\rho(1-H_t(y))}}{1 - e^{-(\mu+\rho)}}, \tag{A.41}$$

which implies that

$$\kappa_t(y_\ell) = \left(\frac{\mu + \rho\kappa_{t-1}(y_\ell)}{\lambda}\right) \frac{\lambda e^{-\lambda}}{1 - e^{-\lambda}} \tag{A.42}$$

and

$$\kappa_t(y_h) = \left(\frac{\mu + \rho\kappa_{t-1}(y_h)}{\lambda}\right) \frac{\lambda}{1 - e^{-\lambda}}. \tag{A.43}$$

Consider the case $t = 1$. Using (A.42), we have $\kappa_2(y_\ell) < \kappa_1(y_\ell)$ if and only if

$$\kappa_2(y_\ell) = \frac{\lambda e^{-\lambda}}{1 - e^{-\lambda}} < \kappa_1(y_\ell) = 1, \tag{A.44}$$

which is true. Using expression (A.43), we have $\kappa_2(y_h) > \kappa_1(y_h)$ if and only if

$$\kappa_2(y_h) = \frac{\lambda}{1 - e^{-\lambda}} > \kappa_1(y_h) = 1, \tag{A.45}$$

which is true. Therefore, we have $\kappa_2(y_\ell) < \kappa_1(y_\ell)$ and $\kappa_2(y_h) > \kappa_1(y_h)$.

Now suppose that $\kappa_t(y_\ell) < \kappa_{t-1}(y_\ell)$ and $\kappa_t(y_h) > \kappa_{t-1}(y_h)$. Expression (A.42) implies that $\kappa_{t+1}(y_\ell) < \kappa_t(y_\ell)$ if and only if

$$\mu + \rho\kappa_t(y_\ell) < \mu + \rho\kappa_{t-1}(y_\ell), \tag{A.46}$$

which is true since we assume that $\kappa_t(y_\ell) < \kappa_{t-1}(y_\ell)$. Next, expression (A.43) implies that $\kappa_{t+1}(y_h) > \kappa_t(y_h)$ if and only if

$$\mu + \rho\kappa_t(y_h) > \mu + \rho\kappa_{t-1}(y_h), \tag{A.47}$$

which is again true since we assume that $\kappa_t(y_h) > \kappa_{t-1}(y_h)$.

Since both $f_{L,t}(y)$ and $f_{R,t}(y)$ are increasing in y, it follows that there exists a unique cut-off y_t^c given by $f_{L,t}(y_t^c) = f_{R,t}(y_t^c)$ such that $\kappa_{t+1}(y_t^c) = \kappa_t(y_t^c)$. It is given by

$$\frac{\lambda e^{-\lambda(1-\Gamma_t(y_t^c))}}{1 - e^{-\lambda}} = \frac{\lambda}{\frac{\mu}{\kappa_t(y_t^c)} + \rho} \tag{A.48}$$

and we have $f_{L,t}(y) > f_{R,t}(y)$ if and only if $y > y_t^c$. So, there exists a unique cut-off y_t^c such that $\kappa_{t+1}(y) > \kappa_t(y)$ if and only if $y > y_t^c$.

We can also prove that the cut-off y_t^c is increasing over time. Consider y_{t+1}^c, which solves:

$$\frac{\lambda e^{-\lambda(1-\Gamma_{t+1}(y_{t+1}^c))}}{1 - e^{-\lambda}} = \frac{\lambda}{\frac{\mu}{\kappa_{t+1}(y_{t+1}^c)} + \rho}. \tag{A.49}$$

First, we have $f_{L,t+1}(y_t^c) < f_{L,t}(y_t^c)$ because $\Gamma_t(y)$ is strictly decreasing in t. Also, $f_{L,t}(y_t^c) = f_{R,t}(y_t^c)$ by definition of y_t^c, so this implies $f_{L,t+1}(y_t^c) < f_{R,t}(y_t^c)$. Given that $\kappa_{t+1}(y_t^c) = \kappa_t(y_t^c)$ by definition of y_t^c, we have $f_{R,t+1}(y_t^c) = f_{R,t}(y_t^c)$. Therefore, we have $f_{L,t+1}(y_t^c) < f_{R,t+1}(y_t^c)$. By definition of y_{t+1}^c, we have $f_{L,t+1}(y) < f_{R,t+1}(y)$ if and only if $y < y_{t+1}^c$, so this implies $y_t^c < y_{t+1}^c$.

(ii) We know that y_t^c is increasing over time and it is bounded above by y_h. Let $y_{\max}^c = \sup\{y_t^c | t = 1, 2, 3, ...\}$. If $y \geq y_{\max}^c$, then $y > y_t^c$ for all t and therefore $\kappa_t(y)$ is always increasing over time. Note that $y_{\max}^c$ solves

$$\frac{\lambda e^{-\lambda(1-\Gamma^*(y_{\max}^c))}}{1 - e^{-\lambda}} = \frac{\lambda}{\frac{\mu}{\kappa^*(y_{\max}^c)} + \rho}\ . \tag{A.50}$$

(iii) Let $y_{\min}^c = \inf\{y_t^c | t = 1, 2, 3, ...\}$, which is y_1^c. If $y < y_{\min}^c$ then $y < y_t^c$ for all t because y_t^c is increasing over time. So if $y < y_{\min}^c$, $\kappa_t(y)$ is always decreasing over time.

Note that y_1^c solves

$$\frac{\lambda e^{-\lambda(1-\Phi(y_1^c))}}{1-e^{-\lambda}} = 1. \tag{A.51}$$

(iv) Suppose that $y \in (y_{\min}^c, y_{\max}^c)$. In period $t = 1$, we have $y > y_1^c$, so $\kappa_2(y) > \kappa_1(y) = 1$. If $y < y_{\max}^c$, there exists a time τ such that $y < y_\tau^c$ and therefore $\kappa_{\tau+1}(y) < \kappa_\tau(y)$. So, the function $\kappa_t(y)$ is initially increasing at $t = 1$ and eventually decreasing. Given that y_t^c is strictly increasing over time, there exists a unique time t_c at which, prior to that time, $y > y_t^c$ and $\kappa_t(y)$ is increasing for all $t < t_c$ and, after that time, $y > y_t^c$ and $\kappa_t(y)$ is decreasing for all $t < t_c$.

Proof of Proposition 8:

(i) Let $\lambda = \mu + \rho$ and let $\hat{H}_t(y) = e^{-\lambda(1-\Gamma_t(y))}$, the distribution of the highest quality among the sellers a buyer meets. Note that $\hat{H}_t(y)$ has a mass point at y_l with probability mass $e^{-\lambda}$ (i.e. the probability a buyer contacts zero sellers) but we can ignore below this because $s_t(y_l) = 0$ for all t. Expression (3.18) is equivalent to

$$s_t(y) = \frac{\int_{y_\ell}^y x\hat{H}_t'(x)dx}{\hat{H}_t(y)}, \tag{A.64}$$

where $\beta = \rho/\lambda$ and

$$\Gamma_t(y) = \beta H_{t-1}(y) + (1-\beta)\Phi(y). \tag{A.65}$$

Similarly, we obtain

$$s_{t+1}(y) = \frac{\int_{y_\ell}^y x\hat{H}_{t+1}'(x)dx}{\hat{H}_{t+1}(y)}. \tag{A.66}$$

We need to show that there exists a unique cut-off quality $y_t^s \in (y_{t+1}^c, y_h)$ such that, if $y > y_t^s$, we have $s_{t+1}(y) > s_t(y)$. Rearranging the above expressions (A.64) and (A.66), $s_{t+1}(y) > s_t(y)$ is equivalent to

$$\int_{y_\ell}^y x\hat{H}_{t+1}'(x)dx \left(\frac{\hat{H}_t(y)}{\hat{H}_{t+1}(y)}\right) > \int_{y_\ell}^y x\hat{H}_t'(x)dx. \tag{A.67}$$

It is easy to see that $\hat{H}_{t+1}(y) < \hat{H}_t(y)$ if and only if $\Gamma_{t+1}(y) < \Gamma_t(y)$, which is true by Proposition 6. Therefore, $\hat{H}_t(y)/\hat{H}_{t+1}(y) > 1$ and it suffices to show that

$$\int_{y_\ell}^y x\left(\hat{H}_{t+1}'(x) - \hat{H}_t'(x)\right)dx \geq 0. \tag{A.68}$$

From (3.17), we have $H_t'(x) = \kappa_{t+1}(x)\Phi'(x)$, and $\hat{H}_t'(x) = H_t'(x)(1-e^{-\lambda})$. Therefore,

$$\hat{H}_t'(x) = \kappa_{t+1}(x)\Phi'(x)(1-e^{-\lambda}). \tag{A.69}$$

Substituting (A.69) into (A.68), we have $s_{t+1}(y) > s_t(y)$ if the following holds:

$$G_t(y) \equiv \int_{y_\ell}^{y} x\Phi'(x)\left(\kappa_{t+2}(x) - \kappa_{t+1}(x)\right) dx \geq 0. \tag{A.70}$$

For $y = y_{t+1}^c$, we have $\kappa_{t+2}(y_{t+1}^c) = \kappa_{t+1}(y_{t+1}^c)$ by definition of y_{t+1}^c and Lemma 7 tells us $\kappa_{t+2}(x) < \kappa_{t+1}(x)$ if and only if $x < y_{t+1}^c$. Therefore, since $\Phi'(x) > 0$ by assumption, we have $G_t(y_{t+1}^c) < 0$. Also, we know that $G_t(y_h) > 0$ because $s_{t+1}(y_h) > s_t(y_h)$ by part (iv). In order to prove that there exists a unique cut-off quality $y_t^s \in (y_{t+1}^c, y_h)$ such that, if $y > y_t^s$, we have $s_{t+1}(y) > s_t(y)$, it therefore suffices to show that $G_t'(y) > 0$ for all $y \in (y_{t+1}^c, y_h)$. For any given $y \in (y_{t+1}^c, y_h)$, we can split $G_t(y)$ as follows:

$$G_t(y) = \int_{y_\ell}^{y_{t+1}^c} x\Phi'(x)\left(\kappa_{t+2}(x) - \kappa_{t+1}(x)\right) dx + \int_{y_{t+1}^c}^{y} x\Phi'(x)\left(\kappa_{t+2}(x) - \kappa_{t+1}(x)\right) dx. \tag{A.71}$$

Differentiating the above, we obtain

$$G_t'(y) = \frac{d}{dy}\int_{y_{t+1}^c}^{y} x\Phi'(x)\left(\kappa_{t+2}(x) - \kappa_{t+1}(x)\right) dx. \tag{A.72}$$

By Lemma 7, we know that $\kappa_{t+2}(x) > \kappa_{t+1}(x)$ if and only if $x > y_{t+1}^c$. Therefore, since $\Phi'(x) > 0$ by assumption, the integrand in (A.72) is always strictly positive over the range (y_{t+1}^c, y). Given that this is true for any $y \in (y_{t+1}^c, y_h)$, we know that the integral in (A.72) above must be increasing in its upper limit y. Therefore, part (i) is proven and the cut-off $y_t^s \in (y_{t+1}^c, y_h)$ is the unique solution to $G_t(y_t^s) = 0$.

(ii) For any given y, define $F_t(x; y) \equiv \frac{\hat{H}_t(x)}{\hat{H}_t(y)}$ where $\hat{H}_t(y) = e^{-\lambda(1-\Gamma_t(y))}$. The function $F_t(; y)$ is a cdf with support $[y_l, y]$ and pdf $f_t(; y)$ given by $f_t(x; y) = \frac{\hat{H}_t'(x)}{\hat{H}_t(y)}$. Using expressions (A.64) and (A.66), we can write $s_t(y) = \int_{y_\ell}^{y} x f_t(x; y) dx$ and $s_{t+1}(y) = \int_{y_\ell}^{y} x f_{t+1}(x; y) dx$. We need to show that if $y < y_t^c$ then $s_{t+1}(y) < s_t(y)$. To prove this, it suffices to show that $F_t(; y)$ strictly first-order stochastically dominates $F_{t+1}(; y)$, i.e. $F_t(x; y) < F_{t+1}(x; y)$ for all $x \in (y_l, y)$. That is, it suffices to show that, for $y < y_t^c$,

$$\frac{\hat{H}_t(x)}{\hat{H}_t(y)} < \frac{\hat{H}_{t+1}(x)}{\hat{H}_{t+1}(y)}, \tag{A.73}$$

or equivalently

$$\frac{\hat{H}_{t+1}(y)}{\hat{H}_t(y)} < \frac{\hat{H}_{t+1}(x)}{\hat{H}_t(x)} \tag{A.74}$$

for any $x \in (y_l, y)$. Since $x < y$, it is sufficient to show that $\frac{\hat{H}_{t+1}(y)}{\hat{H}_t(y)}$ is strictly decreasing in y for any $y < y_t^c$. Recalling that $\hat{H}_t(y) = e^{-\lambda(1-\Gamma_t(y))}$, we have

$$\frac{\hat{H}_{t+1}(y)}{\hat{H}_t(y)} = \frac{e^{-\lambda(1-\Gamma_{t+1}(y))}}{e^{-\lambda(1-\Gamma_t(y))}} = e^{\lambda(\Gamma_{t+1}(y)-\Gamma_t(y))}. \tag{A.75}$$

Differentiating the above, we obtain

$$\frac{d}{dy}\frac{\hat{H}_{t+1}(y)}{\hat{H}_t(y)} = \lambda((\Gamma'_{t+1}(y) - \Gamma'_t(y))e^{\lambda(\Gamma_{t+1}(y)-\Gamma_t(y))}. \tag{A.76}$$

Therefore, we have $\frac{d}{dy}\frac{\hat{H}_{t+1}(y)}{\hat{H}_t(y)} < 0$ if and only if $\Gamma'_{t+1}(y) < \Gamma'_t(y)$. Differentiating expression (A.65) for the distribution $\Gamma_t(y)$, we obtain

$$\Gamma'_t(y) = \beta H'_{t-1}(y) + (1-\beta)\Phi'(y). \tag{A.77}$$

Therefore, $\Gamma'_{t+1}(y) < \Gamma'_t(y)$ if and only if $H'_t(y) < H'_{t-1}(y)$. From (3.17), we have $H'_{t-1}(y) = \kappa_t(y)\Phi'(y)$, so this is true if and only if $\kappa_{t+1}(y)\Phi'(y) < \kappa_t(y)\Phi'(y)$. Since $\Phi'(y) > 0$ by assumption, this is equivalent to $\kappa_{t+1}(y) < \kappa_t(y)$. By Lemma 7, $\kappa_{t+1}(y) < \kappa_t(y)$ if and only if $y < y^c_t$. Thus, $\frac{d}{dy}\frac{\hat{H}_{t+1}(y)}{\hat{H}_t(y)} < 0$ and $s_{t+1}(y) < s_t(y)$ for any $y < y^c_t$.

(iii) Follows from the fact that $s_t(y_h) = (1-e^{-\lambda})\tilde{y}_t$ by Theorem 2, where $\tilde{y}_t$ is the average transaction quality, and $\tilde{y}_t$ is increasing over time by Proposition 6.

(iv) Follows from part (ii). If $y < y^c_{\min} = y^c_1$ where $y^c_1 \in (y_\ell, y_h)$, then the surplus $s_t(y)$ is strictly decreasing at time $t = 1$ by part (ii). By Lemma 7 (i), the cut-off quality y^c_t is increasing over time t. Therefore, $y < y^c_1$ implies that $y < y^c_t$ for any time t, so part (ii) implies the surplus $s_t(y)$ is always strictly decreasing over time.

(v) Suppose that $y \in (y^s_{\min}, y^c_{\max})$. We have $y > y^s_1$, so part (i) implies that $s_t(y)$ is strictly increasing at time $t = 1$. Therefore, since $y^s_t \in (y^c_{t+1}, y_h)$ from the proof of part (i), we must have $y > y^c_2$. Since the sequence y^c_t is strictly increasing over time by Lemma 7 (i) and $y < y^c_{\max}$, there exists $T > 2$ such that $y < y^c_T$. Applying part (ii), the surplus $s_t(y)$ is strictly decreasing at all time periods $t \geq T$. Therefore, the surplus $s_t(y)$ is strictly increasing at first and then eventually strictly decreasing for all $t \geq T$. ■

We also prove the following result, which is mentioned on page 24 in Section 3.

Remark 1. *For unpopular sellers ($\kappa_t(y) < 1$), the surplus function $s_t(y)$ is strictly lower with word of mouth only versus search only, but for popular sellers ($\kappa_t(y) > 1$), $s_t(y)$ is strictly higher with word of mouth only versus search only.*

Proof: Suppose the total expected number of sellers contacted by either search or word of mouth is equal to $\lambda > 0$. If some sellers are contacted *only* through word of mouth, their surplus $s^w_t(y)$ is given by

$$s^w_t(y) = \frac{\int_{y_\ell}^{y} x\hat{H}'_{t,w}(x)dx}{\hat{H}_{t,w}(y)} \tag{A.78}$$

where $\hat{H}_{t,w}(y) = e^{-\lambda(1-H_{t-1}(y))}$. If some sellers are contacted *only* through search, their surplus $s^s_t(y)$ is given by

$$s^s_t(y) = \frac{\int_{y_\ell}^{y} x\hat{H}'_{t,s}(x)dx}{\hat{H}_{t,s}(y)} \tag{A.79}$$

where $\hat{H}_{t,s}(y) = e^{-\lambda(1-\Phi(y))}$. For any given y, define $F_t^w(x;y) \equiv \frac{\hat{H}_{t,w}(x)}{\hat{H}_{t,w}(y)}$ and $F_t^s(x;y) \equiv \frac{\hat{H}_{t.s}(x)}{\hat{H}_{t,s}(y)}$. $F_t^w(.;y)$ and $F_t^s(.;y)$ are cdfs with support $[y_l, y]$ and pdfs $f_t^w(;y)$ and $f_t^s(.;y)$. Using expressions (A.78) and (A.79), $s_t^w(y) = \int_{y_\ell}^y x f_t^w(x;y)dx$ and $s_t^s(y) = \int_{y_\ell}^y x f_t^s(x;y)dx$.

We have $s_t^w(y) > s_t^s(y)$ if $F_t^w(x;y)$ strictly first-order stochastically dominates $F_t^s(x;y)$, i.e. if $F_t^w(x;y) < F_t^s(x;y)$ for all $x \in (y_l, y)$. This is true if and only if

$$\frac{\hat{H}_{t,w}(x)}{\hat{H}_{t,w}(y)} < \frac{\hat{H}_{t.s}(x)}{\hat{H}_{t,s}(y)}, \tag{A.80}$$

which is equivalent to

$$\frac{e^{-\lambda(1-H_{t-1}(x))}}{e^{-\lambda(1-H_{t-1}(y))}} < \frac{e^{-\lambda(1-\Phi(x))}}{e^{-\lambda(1-\Phi(y))}}, \tag{A.81}$$

which is equivalent to $e^{\lambda(H_{t-1}(x)-H_{t-1}(y))} < e^{\lambda(\Phi(x)-\Phi(y))}$, which is true if and only if

$$H_{t-1}(x) - \Phi(x) < H_{t-1}(y) - \Phi(y) \tag{A.82}$$

for all $x \in (y_l, y)$. Since $x < y$, it suffices to show that $H_{t-1}(y) - \Phi(y)$ is strictly increasing in y. Differentiating, this is true if and only if $H'_{t-1}(y) > \Phi'(y)$. From (3.17), we have $H'_{t-1}(y) = \kappa_t(y)\Phi'(y)$, so $H'_{t-1}(y) > \Phi'(y)$ is equivalent to $\kappa_t(y)\Phi'(y) > \Phi'(y)$. Because $\Phi'(y) > 0$ by assumption, this is equivalent to $\kappa_t(y) > 1$. Therefore, we have $F_t^w(x;y) < F_t^s(x;y)$ for all $x \in (y_l, y)$ and thus $s_t^w(y) > s_t^s(y)$ if $\kappa_t(y) > 1$ (i.e. for popular sellers). Conversely, if $\kappa_t(y) < 1$ (i.e. unpopular sellers), the opposite is true because $F_t^w(x;y) > F_t^s(x;y)$ for all $x \in (y_l, y)$ and thus $s_t^w(y) < s_t^s(y)$. ■

Proof of Proposition 13: Let $x(y)$ denote the fraction $\Phi(y)$ of sellers with quality smaller than y, and let $y(x)$ denote the inverse of $x(y)$. Let $z_t(y)$ denote the fraction $H_t(y)$ of transactions made by sellers with quality smaller than y in period t. From (3.15), it follows that $z_1(y)$ is given by

$$z_1(y) = \frac{e^{-\lambda(1-x(y))} - e^{-\lambda}}{1 - e^{-\lambda}}. \tag{A.52}$$

For $t = 1, 2, ...$, it follows from (3.15) that $z_{t+1}(y)$ is given by

$$z_{t+1}(y) = \frac{e^{-\mu(1-x(y))}e^{-\rho(1-z_t(y))} - e^{-(\mu+\rho)}}{1 - e^{-(\mu+\rho)}}. \tag{A.53}$$

The fraction $z_1(x)$ of transactions made by sellers with quality non-greater than $y(x)$ is the solution with respect to z to the equation

$$f(z,x) \equiv z\left(1 - e^{-\lambda}\right) - e^{-\lambda(1-x)} - e^{-\lambda} = 0. \tag{A.54}$$

Consider any $x \in (0,1)$. The function $f(z,x)$ is strictly increasing in z, and such that $f(0,x) < 0$ and $f(1,x) > 0$. Therefore, there exists one and only one $z_1(x)$ that solves the equation $f(z_1(x),x) = 0$ and $z_1(x) \in (0,1)$. The function $g(x) \equiv f(x,x)$ is strictly

concave and such that $g(0) = 0$ and $g(1) = 0$. Therefore, $g(x) = f(x,x) > 0$ for all $x \in (0,1)$. Since $f(x,x) > 0$, $f(z_1(x),x) = 0$, and $f(z,x)$ is strictly increasing in x, it follows that $z_1(x) < x$. For $x = 0$, the unique solution to $f(z,0) = 0$ is $z_1(0) = 0$. For $x = 1$, the unique solution to $f(z,1) = 0$ is $z_1(1) = 1$.

The fraction $z_{t+1}(x)$ of transactions made by sellers with quality non-greater than $y(x)$ is given by

$$f_+(z,x) = \frac{e^{-\mu(1-x)}e^{-\rho(1-z)} - e^{-(\mu+\rho)}}{1 - e^{-(\mu+\rho)}}, \tag{A.55}$$

where z denotes the fraction of transactions made by sellers with quality non-greater than $y(x)$ in period t. Consider any $x \in (0,1)$. For $\rho > 0$, the function $f_+(z,x)$ is strictly increasing and strictly convex in z. The function $f_+(z,x)$ is such that $f_+(0,x) > 0$ and $f_+(1,x) < 1$. Therefore, there exists one and only one $z^*(x)$ that solves the equation $f_+(z,x) = z$ with respect to z for any $x \in (0,1)$. Moreover, $f_+(z,x) \in (z, z^*(x))$ for every $z \in [0, z^*(x))$, and $f_+(z,x) \in (z^*(x), z)$ for every $z \in (z^*(x), 1]$. For $x = 0$, the unique solution to $f_+(z,x) = z$ is $z^*(0) = 0$. For $x = 1$, the relevant solution to $f_+(z,x) = z$ is $z^*(1) = 1$.

It is easy to see that $f_+(x,x) < x$ for any $x \in (0,1)$. Since $f_+(z,x) > z$ for every $z \in [0, z^*(x))$, and $f_+(z,x) < z$ for every $z \in (z^*(x), 1]$, $f_+(x,x) < x$ implies $z^*(x) < x$. In turn, for $\rho > 0$, $z^*(x) < x$ implies

$$\begin{aligned} z^*(x) &= \frac{e^{-\mu(1-x)}e^{-\rho(1-z^*(x))} - e^{-(\mu+\rho)}}{1 - e^{-(\mu+\rho)}} \\ &< \frac{e^{-\mu(1-x)}e^{-\rho(1-x)} - e^{-(\mu+\rho)}}{1 - e^{-(\mu+\rho)}} \\ &= \frac{e^{-\lambda(1-x)} - e^{-\lambda}}{1 - e^{-\lambda}} = z_1(x). \end{aligned} \tag{A.56}$$

Rewriting (A.55) using $\lambda = \mu + \rho$ and $\beta = \rho/\lambda$, we obtain (3.40) in Proposition 13. This is a cdf only if $\frac{dH^*(y)}{dy} \geq 0$, and it is atomless only if $\frac{dH^*(y)}{dy} > 0$. We have that $H^*(y)$ given by (3.40) is strictly increasing if and only if

$$\frac{dH^*(y)}{dy} = \frac{d}{dy}\frac{e^{-\mu(1-\Phi(y))}e^{-\rho(1-H^*(y))} - e^{-(\mu+\rho)}}{1 - e^{-(\mu+\rho)}} > 0. \tag{A.57}$$

Differentiating, this is equivalent to

$$\frac{dH^*(y)}{dy} = \frac{(\mu + \rho\kappa^*(y))\,\Phi'(y)e^{-\mu(1-\Phi(y))}e^{-\rho(1-H^*(y))}}{1 - e^{-(\mu+\rho)}} > 0. \tag{A.58}$$

It is clear that $\frac{dH^*(y)}{dy} > 0$ if and only if $\mu + \rho\kappa^*(y) > 0$ where

$$\kappa^*(y) = \frac{\mu e^{-\mu(1-\Phi(y))}e^{-\rho(1-H^*(y))}}{1 - e^{-\lambda} - \rho e^{-\mu(1-\Phi(y))}e^{-\rho(1-H^*(y))}} \tag{A.59}$$

and therefore

$$\mu + \rho\kappa^*(y) = \frac{\mu(1-e^{-\lambda})}{1-e^{-\lambda}-\rho e^{-\mu(1-\Phi(y))}e^{-\rho(1-H^*(y))}}, \tag{A.60}$$

which is greater than zero if and only if

$$1-e^{-\lambda}-\rho e^{-\mu(1-\Phi(y))}e^{-\rho(1-H^*(y))} > 0. \tag{A.61}$$

The left-hand side of (A.61) is decreasing in y. For $y = y_\ell$, (A.61) since $1-e^{-\lambda}-\rho e^{-\lambda} > 0$ because $\rho \leq \lambda$. For $y = y_h$, (A.61) holds if and only if $1-e^{-\lambda}-\rho > 0$. If $\rho < 1-e^{-\lambda}$, or equivalently if $\beta < \frac{1-e^{-\lambda}}{\lambda}$, we have $\frac{dH^*(y)}{dy} > 0$ for all $y \in [y_\ell, y_h]$, so $y_u = y_h$ and the distribution $H^*(y)$ is atomless. Also, if $\rho < 1-e^{-\lambda}$, then

$$\kappa^*(y_h) = \frac{\mu}{1-e^{-\lambda}-\rho} \tag{A.62}$$

by (A.59). Substituting (A.62) into (A.48), we obtain $y_u = y_h = y^c_{\max}$. ■

Online Appendix: Memory with buyer entry and exit

We consider a version of the model described in Section 2 in which new buyers enter the market in every period, and old buyers exit the market in every period. Initially all buyers are new entrants in period 1. In period $t = 2, 3, ...$ a fraction $\sigma \in [0, 1]$ of buyers that were in the market in period $t-1$ permanently exit the market, and a double continuum of buyers with measure $b\sigma > 0$ per seller enter the market in period $t+1$. In each period, the measure of buyers per seller is b.

With entry and exit of buyers, the distribution of the number of contacts for a buyer depends on how long that buyer has been searching in the market, i.e. the degree of experience of the buyer. We define the level of *experience* of a buyer as the number of periods i the buyer has been in the market.

A buyer who has been in the market for $i < t$ periods has not exited for $i-1$ periods, which occurs with probability $(1-\sigma)^{i-1}$. Therefore, the measure of buyers who have been in the market for $i < t$ periods is equal to the measure $b\sigma$ who were new entrants i periods ago multiplied by the probability $(1-\sigma)^{i-1}$. A buyer who has been in the market for $i = t$ periods is equal to the measure b of buyers at time $t = 1$ multiplied by the probability $(1-\sigma)^{i-1}$ of not exiting for $i-1$ periods. Therefore, if $\sigma < 1$, the measure of buyers who have been in the market for i periods at time t is given by Lemma 9. The distribution of the number of contacts for a buyer i follows from Lemma 1.

Lemma 9: *For a buyer with experience i, the number n_i of the buyer's contacts in period $t = 1, 2, \ldots$ is distributed as a Poisson with coefficient λ_i, where λ_i is given by*

$$\lambda_i = \frac{\mu}{\delta}[1 - (1-\delta)^i]. \tag{2.20}$$

The measure $b_t(i)$ of buyers with experience i at time $t = 1, 2, \ldots$ is given by

$$b_t(i) = \begin{cases} b\sigma(1-\sigma)^{i-1} \text{ if } i < t \\ b(1-\sigma)^{i-1} \text{ if } i = t \end{cases} \tag{2.21}$$

As in the baseline model, the maximized profit for a seller with quality $y \in [y_\ell, y_h]$ is strictly positive (Lemma 2). As in the baseline model, the distribution $F_t(s)$ of surplus offered by sellers does not have any mass points (Lemma 3). Since the surplus distribution $F_t(s)$ does not have any mass points and the $b_t(i)$ buyers with experience i are in contact with a number of sellers distributed as a Poisson with coefficient λ_i, the profit for a seller with quality y that offers the surplus $s \geq 0$ can be written as

$$V_t(y, s) = \left[\sum\nolimits_{i=1}^{t} b_t(i)\lambda_i e^{-\lambda_i(1-F_t(s))}\right](y - s). \tag{2.22}$$

Note that $\sum_{i=1}^{t} b_t(i)\lambda_i e^{-\lambda_i(1-F_t(s))}$ is the sum of the measure $b_t(i)\lambda_i e^{-\lambda_i(1-F_t(s))}$ of buyers with experience i such that the seller offers the highest surplus for those buyers, i.e. it is the sum of trade volume across buyer types.

As in the baseline model, the support of the surplus distribution $F_t(s)$ is an interval $[s_{\ell,t}, s_{h,t}]$, with $s_{\ell,t} = 0$ (Lemma 4). Moreover, the surplus offered by a seller is a strictly increasing function $s_t(y)$ of the quality y of the variety carried by the seller (Lemma 5). Therefore, as in the baseline model, the fraction $F_t(s_t(y))$ of sellers offering surplus smaller than $s_t(y)$ is equal to the fraction $\Phi(y)$ of sellers with quality smaller than y. That is,

$$F_t(s_t(y)) = \Phi(y). \tag{2.23}$$

For a seller producing quality y, the seller's problem at date t is

$$V_t(y) = \max_{s\in[0,y]} \left[\sum_{i=1}^{t} b_t(i)\lambda_i e^{-\lambda_i(1-F_t(s))}\right](y-s). \tag{2.24}$$

The necessary condition for the optimality of the surplus $s_t(y)$ offered by a seller of quality y in period t is

$$\left[\sum\nolimits_{i=1}^{t} b_t(i)\lambda_i^2 e^{-\lambda_i(1-F_t(s_t(y)))}F_t'(s_t(y))\right](y-s_t(y)) - \sum\nolimits_{i=1}^{t} b_t(i)\lambda_i e^{-\lambda_i(1-F_t(s_t(y)))} = 0. \tag{2.25}$$

The left-hand side of (2.25) is the seller's marginal benefit from offering an additional unit of surplus to its buyers, which is given by the increase in the seller's volume of trade multiplied by the seller's profit per trade. The right-hand side of (2.25) is the seller's marginal cost from offering an additional unit of surplus to its buyers, which is given by the seller's volume of trade.

Using (2.23) to substitute $F_t(s_t(y))$ with $\Phi(y)$ and $F_t'(s_t(y))$ with $\Phi'(y)/s_t'(y)$, we can rewrite the optimality condition (2.25) as

$$s_t'(y) = \Phi'(y)\Lambda_t(y)(y - s_t(y)) \tag{2.26}$$

where $\Lambda_t(y)$ is given by

$$\Lambda_t(y) = \frac{\sum_{i=1}^{t} b_t(i)\lambda_i^2 e^{-\lambda_i(1-F_t(s_t(y)))}F_t'(s_t(y))}{\sum_{i=1}^{t} b_t(i)\lambda_i e^{-\lambda_i(1-F_t(s_t(y)))}}. \tag{2.27}$$

The surplus function $s_t(y)$ must satisfy the differential equation (2.26), together with the boundary condition $s_t(y_\ell) = 0$. The differential equation has the same structure as in the baseline model, except that $\Lambda_t(y)$ takes the place of λ_t. It is easy to see that $\Lambda_t(y)$ is simply a weighted average of the Poisson coefficient λ_i for buyers with different experience levels.

The unique candidate equilibrium of the product market is given by the solution $s_t(y)$ of the differential equation (2.26)) together with the boundary condition $s_t(y) = 0$. Using the same argument used for the baseline model, it is easy to verify that the candidate equilibrium is indeed an equilibrium.

Theorem 4: (Existence and uniqueness of equilibrium with entry and exit) *An equilibrium exists and is unique. For $t = 1, 2, ...,$ equilibrium is given by a surplus function $s_t(y)$ that satisfies the differential equation (2.26)) together with the boundary condition $s_t(y_\ell) = 0$. The solution to this differential equation is*

$$s_t(y) = \left(\int_{y_\ell}^{y} \Lambda_t(x) e^{-(B_t(y) - B_t(x))} x \Phi'(x) dx \right), \tag{2.28}$$

where

$$B_t(y) = \int_{y_\ell}^{y} \Lambda_t(z) \Phi'(z) dz \tag{2.29}$$

and

$$\Lambda_t(y) = \sum_{i=1}^{t} \tilde{\omega}_{i,t}(y) \lambda_i \tag{2.30}$$

where

$$\omega_{i,t}(y) = \frac{b_t(i) \lambda_i e^{-\lambda_i (1 - \Phi(y))}}{\sum_{j=1}^{t} b_t(j) \lambda_j e^{-\lambda_j (1 - \Phi(y))}}. \tag{2.31}$$

Associated with the equilibrium surplus function $s_t(y)$, there are a surplus distribution $F_t(s)$, a price distribution $G_t(p)$, and a transaction distribution $H_t(y)$. The surplus distribution $F_t(s)$ is

$$F_t(s) = \Phi(y_t(s)), \tag{2.32}$$

where $y_t(s)$ is the inverse of the surplus function $s_t(y)$. The price distribution $G_t(p)$ is

$$G_t(p) = \int_{y : p_t(y) \leq p} \Phi'(y) dy, \tag{2.33}$$

where $p_t(y)$ is the pricing function $y - s_t(y)$. The transaction distribution $H_t(y)$ is

$$H_t(y) = \frac{\sum_{i=1}^{t} b_t(i) \left(e^{-\lambda_i (1 - \Phi(y))} - e^{-\lambda_i} \right)}{\sum_{j=1}^{t} b_t(j) \left(1 - e^{-\lambda_j} \right)}, \tag{2.34}$$

where the numerator in (2.34) is the quantity of the good traded by sellers with quality less than y, and the denominator in (2.34) is the quantity of the good traded by all sellers.

4.1 Market dynamics

As in the baseline model, the market becomes more competitive over time because the fraction of more experienced buyers increases over time. To formalize this intuition, we first consider the dynamics of the distribution of buyer experience.

Lemma 10: *The distribution of buyer experience i is the following:*

1. *If* $\sigma \in (0,1)$*, the share of buyers with experience level* i *at time* t *is*

$$\pi_{i,t} = \begin{cases} \sigma(1-\sigma)^{i-1} \text{ if } i < t \\ (1-\sigma)^{i-1} \text{ if } i = t \end{cases} \tag{2.35}$$

   where $\sum_{i=1}^{t} \pi_{i,t} = 1$ *and* $\pi_{i,t}$ *is decreasing in* i*.*

2. *If* $\sigma = 0$*, then* $\pi_{t,t} = 1$ *and* $\pi_{i,t} = 0$ *for* $i < t$*.*

3. *If* $\sigma = 1$ *then* $\pi_{1,t} = 1$ *and* $\pi_{t,t} = 0$ *for* $i > 1$*.*

The average buyer experience at time t is $\tilde{\imath}_t \equiv \sum_{i=1}^{t} \pi_{i,t} i$. Lemma 11 establishes that buyers become more experienced over time.

Lemma 11: *For any* $\sigma \in (0,1)$*, buyers become more experienced over time. That is, the distribution* $\pi_{i,t+1}$ *of buyer experience at time* $t+1$ *first-order stochastically dominates the distribution* $\pi_{i,t}$ *of buyer experience at time* t*. The average buyer experience* $\tilde{\imath}_t$ *is strictly increasing in* t*.*

As buyers become more experienced over time, there is a larger number of sellers they previously contacted which they may potentially remember. As a result, the average number of contacts increases over time. With entry and exit, the relevant measure of the average number is the weighted average $\Lambda_t(y) = \sum_{i=1}^{t} \tilde{\omega}_{i,t}(y)\lambda_i$ where the weights $\tilde{\omega}_{i,t}(y)$ are given by (2.31). Lemma 9 shows that this weighted average number of contacts is increasing over time.

Lemma 12: *The weighted average number of contacts* $\Lambda_t(y)$ *is strictly increasing in* t*.*

Following the same arguments as in Proposition 1, it is easy to show that, since $\Lambda_t(y)$ is strictly increasing in t, the surplus function $s_t(y)$ is strictly increasing over time. Since $s_t(y)$ is strictly increasing in t, it follows from (2.32) that $F_t(s)$ is strictly decreasing in t. Since $s_t(y)$ is strictly increasing in t, the price function $p_t(y) = y - s_t(y)$ is strictly decreasing over time. Therefore, it follows from (2.33) that $G_t(p)$ is strictly increasing in t. Following the same arguments as in Propositions 2 and 3, it is also easy to see that $H_t(y)$ is strictly decreasing in t and, hence, trade concentration is increasing over time. These observations are summarized in the following proposition.

Proposition 16: (Market dynamics with entry and exit) *Let* $\sigma \in [0,1)$ *and* $\delta \in [0,1)$*. For any given* $y < y_\ell$*, the surplus* $s_t(y)$ *is strictly increasing in* t*, and the price* $p_t(y)$ *is strictly decreasing in* t*. The surplus distribution* $F_t(s)$ *is strictly increasing in* t *in the sense of first-order stochastic dominance, the price distribution* $G_t(p)$ *is strictly decreasing in* t *in the sense of first-order stochastic dominance, and the transaction distribution* $H_t(y)$ *is strictly increasing in* t *in the sense of first-order stochastic dominance.*

If the buyers' exit probability σ is equal to 1, the distribution of buyers across experience levels is degenerate, in the sense that $b_t(i) = b$ and $b_t(i) = 0$ for $i = 2, 3, \ldots t+1$. It then follows from (2.27) that $\Lambda_{t+1}(y) - \Lambda_t(y) = 0$ and $\Lambda_t(y) = \mu$ for $t = 1, 2, \ldots$. Hence, if

$\sigma = 1$, the surplus function $s_t(y)$, the price function $p_t(y)$, the surplus distribution $F_t(s)$, the price distribution $G_t(p)$, and the transaction distribution $H_t(y)$ remain constant over time. Intuitively, if $\sigma = 1$, buyers cannot accumulate any experience and, hence, the extent of competitiveness of the product market remains constant over time and so do all the equilibrium outcomes. The same conclusions apply to the case in which buyers have no memory ($\delta = 1$), since in this case buyers do accumulate experience but experience is worthless.

We now want to compare the properties of equilibrium if the buyers' exit probability is lower (σ_1) or higher (σ_2), given that buyers have some memory. In period 1, $\Lambda_1(y) = \mu$ irrespective of σ and, hence, the equilibrium is the same whether the buyers' exit probability is lower (σ_1) or higher (σ_2). Formally, in period 1, $F_1(s|\sigma_1) = F_1(s|\sigma_2)$, $G_1(p|\sigma_1) = G_1(p|\sigma_2)$ and $H_1(y|\sigma_1) = H_1(y|\sigma_2)$. In any period $t \geq 2$, $\Lambda_t(y)$ is strictly higher if the buyer's exit probability is lower (σ_1) than higher (σ_2), since the distribution of buyers is tilted towards buyers with more experience when the buyers' exit probability is lower. For this reason, in any period $t \geq 2$, the market is more competitive if the buyer's exit probability is lower. In turn, this implies that, in any period $t \geq 2$, the surplus offered by sellers is higher, the prices posted by sellers are lower, and the transactions are shifted toward higher-quality sellers if the buyers' exit probability is σ_1 rather than σ_2. Formally, for $t \geq 2$, $F_t(s|\sigma_1) < F_t(s|\sigma_2)$, $G_t(p|\sigma_1) > G_t(p|\sigma_2)$ and $H_t(y|\sigma_1) < H_t(y|\sigma_2)$.

The above observations are summarized in the following proposition.

Proposition 17: (Exit). *Take any $\delta \in [0, 1)$ and any σ_1 and σ_2, with $0 \leq \sigma_1 < \sigma_2 \leq 1$.*

(i) *In period 1, equilibrium surplus, prices and transactions are the same with lower exit, in the sense that $F_1(s|\sigma_1) = F_1(s|\sigma_2)$, $G_1(p|\sigma_1) = G_1(p|\sigma_2)$ and $H_1(y|\sigma_1) = H_1(y|\sigma_2)$.*

(ii) *In any period $t \geq 2$, equilibrium surplus is higher, prices are lower, and transactions are better with lower exit, in the sense that $F_t(s|\sigma_1) < F_t(s|\sigma_2)$, $G_t(p|\sigma_1) > G_t(p|\sigma_2)$ and $H_t(y|\sigma_1) < H_t(y|\sigma_2)$.*

(iii) *Let $\sigma_2 = 1$. With full exit, equilibrium surplus, prices and transactions are constant over time, in the sense that $F_t(s|\sigma_2) = F_1(s|\sigma_2)$, $G_t(p|\sigma_2) = G_1(p|\sigma_2)$, and $H_t(y|\sigma_2) = H_1(y|\sigma_2)$.*

4.2 Steady state

We now consider the stationary limit. The limiting behavior of the product market is determined by the limiting behavior of the weighted average number of contacts $\Lambda_t(y)$. In the limit as $t \to \infty$,

$$\lim_{t\to\infty} \Lambda_t(y) = \lim_{t\to\infty} \frac{\sum_{i=1}^{t} b\sigma(1-\sigma)^{i-1}\lambda_i^2 F_t'(s_t(y))e^{-\lambda_i(1-F_t(s_t(y)))}}{\sum_{i=1}^{t} b\sigma(1-\sigma)^{i-1}\lambda_i e^{-\lambda_i(1-F_t(s_t(y)))}}. \tag{2.36}$$

Since $\Lambda_t(y)$ is increasing in t, $\Lambda_t(y)$ converges to some finite $\Lambda^*(y)$ if and only if $\Lambda_t(y)$ is bounded from above. For any $\sigma \in (0,1]$, the numerator is the sum of products between a term that declines exponentially, $(1-\sigma)^{i-1}$, and a term λ_i^2 that is bounded above by $(\mu i)^2$. Therefore, the numerator in (2.36) is bounded from above. Since the denominator in (2.36) is strictly increasing in t and the numerator in (2.36) is bounded from above, $\Lambda_t(y)$ is bounded from above and, hence, $\Lambda_t(y)$ converges to some finite $\Lambda^*(y)$. For $\sigma = 1$ and any $\delta \in (0,1]$, $\Lambda_t(y)$ is a weighted average of λ_i for $i = 1, \dots t$. Since λ_i is bounded above by μ/δ, it follows that (2.36) is bounded from above, and, hence, $\Lambda_t(y)$ converges to some finite $\Lambda^*(y)$.

From the above observations, it follows that the limiting behavior of the product market remains imperfectly competitive as long as either buyers have imperfect memory, in the sense that $\delta \in (0,1]$, or buyers have a strictly positive probability of exiting, in the sense that $\sigma \in (0,1]$. In fact, if $\delta \in (0,1]$ or $\sigma \in (0,1]$, $\Lambda_t(y)$ converges to some finite $\Lambda^*(y)$ and the equilibrium associated with a finite $\Lambda^*(y)$ is such that sellers enjoy a strictly positive profit, prices are strictly greater than marginal cost, and all sellers, including those with lower quality, trade the good. These findings are intuitive. If buyers have imperfect memory, there is always a positive fraction of buyers that are captive. Hence, sellers' profits are always bounded away from zero and the market remains imperfectly competitive. Similarly, if buyers have a strictly positive probability of exit, there is always a positive fraction of buyers who have just entered the market and have limited contacts with sellers. Hence, sellers' profits are always bounded away from zero and the market remains imperfectly competitive.

Up to this point, we focused on the dynamics of the product market. First, we showed that, as long as buyers have some memory and as long as buyers have a probability of exiting that is less than 1, the market becomes progressively more competitive and, for this reason, the surplus offered by sellers increases over time, the prices posted by sellers decrease over time, and the distribution of transactions moves from lower to higher quality sellers. Second, we showed that, as long as buyers do not have perfect memory or as long as buyers have a strictly positive probability of exiting, the market does not become perfectly competitive.

The version of the model with entry and exit of buyers also allows us to study the effect of *experience* on buyers' outcomes independently from the dynamics of the product market. To analyze the effect of experience, suppose that $\delta \in (0,1)$ and $\sigma \in (0,1)$ and that the market has reached its stationary limit.[13]

[13]We focus on the stationary limit because, in such limit, the difference in outcomes across buyers with different experience i are the same as the change in outcomes for an individual buyer j as its experience i increases. In an arbitrary period t, the differences in outcomes across buyers with different experience i are qualitatively identical to those in the stationary limit. In an arbitrary period t, however, the difference in outcomes across buyers with different experience i are not the same as the change in outcomes for an individual buyer j as its experience i grows, since the latter would have to include changes in market-wide outcomes between t and $t+1$.

In this stationary limit, the distribution of buyer experience converges to a steady state distribution. As $t \to \infty$, we have $\pi_i^* = \sigma(1-\sigma)^{i-1}$ for all i, and the average buyer experience is just $\tilde{\imath}^* = 1/\sigma$. For $\delta > 0$, the average number of contacts for buyers with experience i is $\lambda_i = \frac{\mu}{\delta}[1-(1-\delta)^i]$, and the steady state average number of contacts per buyer, λ^*, is given by

$$\lambda^* = \frac{\mu}{\delta}\left[1 - \frac{\sigma(1-\delta)}{\delta + \sigma(1-\delta)}\right]. \tag{2.37}$$

It is clear that $\lambda^* < \mu/\delta$ whenever $\delta < 1$ and $\sigma > 0$. This makes sense because buyers have limited experience and therefore smaller choice sets compared to the model with memory and infinitely-lived buyers in Section 2.

Buyers' experience affects the quality of the goods they purchase and the value of the surplus they receive. Buyers with experience i are in contact with n_i sellers, where n_i is distributed as a Poisson with coefficient λ_i. Among the buyers with experience i that purchase the good, the distribution of surplus across transactions, $\tilde{F}_i(s)$, is given by

$$\begin{aligned} \tilde{F}_i(s) &= \left[\sum_{k=1}^{\infty} \frac{e^{-\lambda_i}\lambda_i^k}{k!} F^*(s)^k\right] \Big/ \left[1 - e^{-\lambda_i}\right] \\ &= \frac{e^{-\lambda_i(1-F^*(s))} - e^{-\lambda_i}}{1 - e^{-\lambda_i}}. \end{aligned} \tag{2.38}$$

by similar reasoning to the derivation of expression (2.6) for $H_t(y)$. Using the same results as in the proof of Proposition 2, it is easy to show that $\lambda_{i+1} > \lambda_i$ implies $F_{i+1}(s) < F_i(s)$. As a result, the average surplus for a transaction for a buyer with experience i is strictly increasing with respect to the buyer's experience. Similarly, the average transaction quality for a buyer with experience i is strictly increasing with respect to the buyer's experience.

Proposition 18: *On average, more experienced buyers purchase higher quality goods and receive higher surplus than less experienced buyers.*

It is useful to point out that, while the buyers' average surplus increases with experience, the surplus of a specific buyer may decrease, remain constant, or increase depending on how many contacts the buyer has, how many contacts the buyer loses, and how many new contacts the buyer adds. For instance, if some buyer j has $n_{i,j}$ contacts, loses the best of its contacts, and does not acquire any new contacts, its surplus falls. If buyer j has $n_{i,j}$ contacts, does not lose any of its contacts, and does not acquire any new contacts, its surplus remains unchanged. If buyer j reaches a seller that is better than any of its previous contacts, the buyer's surplus increases. Moreover, all of these three events have a strictly positive probability of occurring.

Proofs for Online Appendix

Proof of Theorem 4: We have already established existence and uniqueness of equilibrium. Rewriting the differential equation (2.26), the function $s_t(y)$ solves the following differential equation:

$$s_t'(y) + \Lambda_t(y)\Phi'(y)s_t(y) = \Lambda_t(y)y\Phi'(y) \tag{A.17}$$

where $s_t(y_\ell) = 0$. For any given y, there exists a unique solution $s_t(y)$ which satisfies this differential equation. The general form of the solution is

$$s_t(y) = e^{-\int_{y_\ell}^{y} \Lambda_t(z)\Phi'(z)dz} \left(\int_{y_\ell}^{y} e^{\int_{y_\ell}^{x} \Lambda_t(z)\Phi'(z)dz} \Lambda_t(x)x\Phi'(x)dx + c \right), \tag{A.18}$$

where c is a constant. Given that $s_t(y_\ell) = 0$, we have $c = 0$. Therefore,

$$s_t(y) = \left(\int_{y_\ell}^{y} e^{-(B_t(y) - B_t(x))} \Lambda_t(x)x\Phi'(x)dx \right) \tag{A.19}$$

where $B_t(y) = \int_{y_\ell}^{y} \Lambda_t(z)\Phi'(z)dz$. ■

Proof of Lemma 10: Dividing the measure of buyers who have been in the market for i years by the total measure of buyers b, we obtain (2.35). Next, we have

$$\sum_{i=1}^{t} \pi_{i,t} = \sum_{i=1}^{t-1} \sigma(1-\sigma)^{i-1} + (1-\sigma)^{t-1} \tag{A.20}$$

where $\sum_{i=1}^{t-1} \sigma(1-\sigma)^{i-1} = \sigma \sum_{i=0}^{t-2} (1-\sigma)^i$, which is equal to $\sigma \left(\frac{1-(1-\sigma)^{t-1}}{1-(1-\sigma)} \right) = 1 - (1-\sigma)^{t-1}$ for $\sigma < 1$. Therefore, $\sum_{i=1}^{t} \pi_{i,t} = 1 - (1-\sigma)^{t-1} + (1-\sigma)^{t-1} = 1$. If $\sigma = 0$, then $\pi_{t,t} = 1$ follows from Part 1. If $\sigma = 1$, all buyers are new entrants in each period so $\pi_{1,t} = 1$. ■

Proof of Lemma 11: At time $t+1$ and time t, the share of buyers with experience level i is the same for all $i < t$ because $\pi_{i,t+1} = \pi_{i,t} = \sigma(1-\sigma)^{i-1}$. However, for $i = t$, we have $\pi_{i,t+1} = \sigma(1-\sigma)^{i-1}$ and $\pi_{i,t} = (1-\sigma)^{i-1}$, so $\pi_{i,t+1} < \pi_{i,t}$ for $\sigma < 1$. Therefore, $\sum_{i=1}^{t} \pi_{i,t+1} < \sum_{i=1}^{t} \pi_{i,t}$ and this is true for any $t \geq 1$. Thus $\tilde{\imath}_t$ is strictly increasing in t. ■

Proof of Lemma 12: It is sufficient to examine the difference between $\Lambda_{t+1}(y)$ and $\Lambda_t(y)$, which can be written as

$$\Lambda_{t+1}(y) - \Lambda_t(y) = \left[\sum\nolimits_{i=1}^{t} \left(\omega_{i,t+1}(y) - \omega_{i,t}(y) \right) \lambda_i \right] + \omega_{t+1,t+1}(y)\lambda_{t+1}, \tag{A.21}$$

where

$$\omega_{i,t}(y) = \frac{b_t(i)\lambda_i e^{-\lambda_i(1-\Phi(y))}}{\sum_{j=1}^{t} b_t(j)\lambda_j e^{-\lambda_j(1-\Phi(y))}} \tag{A.22}$$

and

$$\omega_{i,t+1}(y) = \frac{b_t(i)\lambda_i e^{-\lambda_i(1-\Phi(y))}}{\sum_{j=1}^{t+1} b_t(j)\lambda_j e^{-\lambda_j(1-\Phi(y))}} \tag{A.23}$$

If buyers have some memory, in the sense that $\delta \in [0,1)$, Lemma 9 tells us that the average number λ_i of contacts for a buyer is strictly increasing in the buyer's experience i. If buyers have a positive probability of staying in the market for multiple periods, in the sense that $\sigma \in [0,1)$, the measure of buyers $b(i)$ with experience i is strictly positive, and so are $\omega_{i,t}(y)$ and $\omega_{i,t+1}(y)$. Using these observations yields

$$\begin{aligned}
\Lambda_{t+1}(y) - \Lambda_t(y) &= \left[\sum_{i=1}^{t} \left(\omega_{i,t+1}(y) - \omega_{i,t}(y)\right)\lambda_i\right] + \omega_{t+1,t+1}(y)\lambda_{t+1} \\
&> \left[\sum_{i=1}^{t} \left(\omega_{i,t+1}(y) - \omega_{i,t}(y)\right)\lambda_{t+1}\right] + \omega_{t+1,t+1}(y)\lambda_{t+1} \\
&= \left[\sum_{i=1}^{t} \left(\omega_{i,t+1}(y) - \omega_{i,t}(y)\right) + \omega_{t+1,t+1}(y)\right]\lambda_{t+1} = 0.
\end{aligned} \tag{A.24}$$

Specifically, the second line in (A.24) makes use of the fact that $b(i) > 0$ for $i = 1, 2, ...t+1$ and, hence, (A.22) and (A.23) imply that $\omega_{i,t+1} - \omega_{i,t} < 0$ for $i = 1, 2, \ldots t$. Moreover, the second line uses the fact that $\lambda_{t+1} > \lambda_i$ for $i = 1, 2, \ldots t$. The third line uses the fact that $\omega_{i,t}(y)$ and $\omega_{i,t+1}(y)$ are weights, in the sense that $\sum_{i=1}^{t} \omega_{i,t}(y) = 1$ and $\sum_{i=1}^{t+1} \omega_{i,t+1}(y) = 1$ and, hence, the term in square brackets multiplying λ_{t+1} is equal to 0. ■