Consumer Choice, Inflation, and Welfare*

Ayushi Bajaj† and Sephorah Mangin‡

June 10, 2020

Abstract

We introduce consumer choice into a search-theoretic model of monetary exchange and examine its implications for efficiency and the welfare cost of inflation. Consumers can simultaneously meet a number of different sellers and *choose* a seller with whom to trade. Consumer choice is influenced by seller-specific random utility ("quality") shocks that are private information for the buyer. In competitive search equilibrium, there may be ranges of underconsumption and overconsumption, and there may be either under-entry or over-entry of sellers. The Friedman rule does not generally deliver the efficient allocation. We calibrate the model and quantitatively evaluate the significance of consumer choice for the welfare cost of inflation. We find that the welfare cost of increasing inflation from 0% to 10% is more than twice as high when we incorporate choice: 1.5% of consumption versus 0.6% without choice.

JEL codes: D82, E31, E40, E50, E52

Keywords: Money; Inflation; Private information; Competitive search

---

*We are particularly grateful to Guillaume Rocheteau and Randy Wright for especially useful feedback. We thank seminar participants at U.C. Irvine and conference participants at the Bank of Canada Summer Workshop on Money, Banking, and Payments for useful comments.

†Department of Economics, Monash University. Email: ayushi.bajaj@monash.edu.

‡Corresponding author. Research School of Economics, Australian National University. Email: sephorah.mangin@anu.edu.au.

1 Introduction

When consumers purchase goods, they typically choose from a number of goods that are available simultaneously from a range of sellers. Choice is often idiosyncratic: different consumers may make different choices when faced with the same range of goods. Search-theoretic models have become the standard way of modelling the microfoundations of monetary exchange, as surveyed in Lagos, Rocheteau, and Wright (2017), but meetings are typically one-on-one or bilateral in these models. Each buyer (consumer) meets at most one seller (firm) during a single period of time and can either purchase from that seller or wait until the next period to trade. In that sense, there is a limited role for what we call *consumer choice.*

This paper shows that consumer choice is an important feature of monetary exchange. One application of search-theoretic models of money, such as Lagos and Wright (2005), is estimating the welfare cost of inflation. To show that choice matters, we ask the question: how does consumer choice affect efficiency and the welfare cost of inflation? To answer this question, we develop a search-theoretic model of money that features consumer choice. In particular, we build on the framework of Rocheteau and Wright (2005), hereafter denoted RW, because it shares the convenience of the Lagos and Wright (2005) structure – with both centralized and decentralized markets – and it also features endogenous seller entry, which will be a key driver of choice.

We focus on *competitive search equilibrium.* Buyers and sellers choose to enter submarkets in which terms of trade, or contracts, are posted by market makers. After entering a submarket, buyers and sellers commit to trading at the terms posted in that submarket. Within each submarket, there are search frictions that govern how buyers and sellers meet. Competitive search is natural in the environment we consider since buyers can meet multiple sellers within a meeting. At the same time, it is a natural benchmark for welfare analysis since competitive search is often used to decentralize the constrained efficient allocation in search-theoretic environments, as discussed in Wright, Kircher, Julien, and Guerrieri (2020). Moreover, since the cost of inflation is generally much lower when prices are determined by competitive search instead of bargaining, our estimate of the welfare cost of inflation with consumer choice is conservative and can be interpreted as a lower bound.

To allow for consumer choice, we model search frictions within submarkets using a random meeting technology that features *many-on-one* or *multilateral* meetings.

In any given period of time, a buyer may simultaneously "meet" a number of sellers but can choose to trade with only one. After a meeting takes place, the buyer draws an i.i.d. *utility shock* for each of the sellers they meet and then chooses to purchase from the seller that maximizes their net utility. Sellers cannot observe buyers' utility shocks; they are private information for the buyer. We sometimes refer to the realization of a shock as the good's *quality*, but it is really perceived quality (or "suitability") since it is an idiosyncratic utility or "taste" shock. What is important is that this seller-specific idiosyncratic shock is what determines each buyer's choice of seller.

Our framework differs from RW in two main respects. First, buyers' utility from consumption depends on both the *quantity* of goods consumed and their (idiosyncratic) utility or *quality*, which is private information for the buyer. Second, consumers have the opportunity to *choose* which goods to purchase among those offered by a number of different sellers. Together, these features imply that the expected match surplus depends directly on the seller-buyer ratio. A larger number of sellers per buyer means that buyers have both a greater meeting probability *plus* greater choice among sellers, which increases the expected quality of the chosen goods.

We focus on incentive-compatible direct revelation mechanisms that induce buyers to reveal their private information to their chosen seller. A mechanism is a list of contracts specifying the quantity traded and payment in real dollars for any given realization of the buyer's utility shock. In equilibrium, there is only one active submarket and sellers offer the same non-linear price schedule. Equilibria may be either *full trade* (all meetings result in trade) or *partial trade* (not all meetings result in trade). Buyers may spend all of their money, some of their money, or none.

In terms of efficiency, there are two margins: an *extensive margin* (seller entry) and an *intensive margin* (quantity traded). With consumer choice, the extensive margin has two components since seller entry directly affects both the number of trades *and* the size of the expected match surplus. We find there may be inefficiencies on both the intensive and extensive margins. In particular, outside the Friedman rule, there are various possibilities for ranges of underconsumption and overconsumption relative to the efficient quantity, and there may be either under-entry or over-entry of sellers.

The Friedman rule does not generally deliver efficiency. Only when the equilibrium is full-trade – which rarely occurs – does the Friedman rule achieve efficiency. In general, the Friedman rule does not deliver efficiency along *either* the extensive or intensive margin. First, there is underconsumption of all goods at the Friedman rule.

Second, there may be either under-entry, over-entry, or efficient entry of sellers at the Friedman rule. These results contrast with those found in RW. In their environment, there may be underconsumption and under-entry of sellers in competitive search equilibrium, but the Friedman rule corrects both types of inefficiency.

After presenting our key analytic results, we use our model to quantify the effect of consumer choice on the welfare cost of inflation. We calibrate the model to match data from Lucas and Nicolini (2015) on money demand in the U.S. from 1915-2008. In order to isolate the effect of consumer choice, we also calibrate an alternative version of the model in which we shut down consumer choice. We find that choice matters: the cost of inflation is significantly higher with consumer choice than without choice. We estimate that the welfare cost of increasing inflation from 0% to 10% is more than twice as high when we allow for consumer choice: 1.45% of consumption compared to 0.60% without choice. The welfare cost of moving from the Friedman rule to 10% inflation is also more than twice as high: 1.65% of consumption versus 0.78%.

The fact that consumer choice leads to a significantly higher welfare cost of inflation is due primarily to two factors: *(i)* the greater effect of inflation on seller entry; and *(ii)* the stronger implications for welfare of any given decline in the seller-buyer ratio. When inflation rises from 0% to 10%, we find that the seller-buyer ratio falls by a much higher amount with choice than without choice. Moreover, due to changes in both the average quantity traded and the average quality of chosen goods, the average match surplus decreases more dramatically with choice than without choice (where there is no change in average quality). As a result, there is a much greater decrease in the total surplus created in the decentralized market, which falls by more than twice as much when we allow the possibility of choice.

In our baseline calibration, there is underconsumption of all goods and under-entry of sellers both with and without choice. At the Friedman rule, there is still underconsumption of all goods both with and without choice. However, there is *over-entry* of sellers with choice and under-entry without choice. We calculate the welfare loss at the Friedman rule compared to the efficient allocation, both with and without choice. We estimate that the welfare loss at the Friedman rule is 0.47% of consumption without consumer choice, but only 0.16% with choice. Intuitively, the fact that there is over-entry of sellers with choice leads to a slightly higher average quality at the Friedman rule compared to the efficient outcome, which offsets the reduction in the average quantity traded and alleviates the extent of the welfare loss.

Outline of the paper. Section 2 discusses the relationship between our paper and the existing literature. Section 3 describes the basic structure of the model. Section 4 derives some results regarding the distribution of chosen goods. Section 5 solves the planner's problem. Section 6 describes competitive search equilibrium and establishes existence and uniqueness. Section 7 presents our key analytic results. Section 8 describes our calibration of the model and presents our estimates of the welfare cost of inflation. Section 9 concludes. All proofs are in the Appendix.

2 Related literature

There is a large literature on the welfare cost of inflation. Cooley and Hansen (1989) estimate that the cost of 10% inflation is less than 0.5% of consumption, while Lucas (2000) estimates that it is less than 1%. Lagos and Wright (2005) find that the cost of 10% inflation is between 3% and 5% of consumption in a monetary model with search and bargaining. In competitive search equilibrium, the cost of inflation is typically much lower than under bargaining. Rocheteau and Wright (2009) estimate that the cost of 10% inflation is between 0.67% and 1.1% of consumption.[1] More recently, Bethune, Choi, and Wright (2020) develop a monetary model with both directed search (by informed buyers) and random search (by uninformed buyers). They obtain a relatively low estimate for the cost of inflation – around 1% or less – by identifying a positive market-composition effect of inflation.

As discussed, our model builds on the environment in RW, which features seller entry but does not feature consumer choice. In RW, the focus is on comparing different market structures, while our paper examines the effect of consumer choice on efficiency and the welfare cost of inflation in competitive search equilibrium. Our paper is related to the broad literature on competitive search surveyed in Wright et al. (2020). In particular, we contribute to the literature on competitive search and private information, including Faig and Jerez (2005), Guerrieri (2008), Guerrieri, Shimer, and Wright (2010), Moen and Rosen (2011), and Davoodalhosseini (2019). In our environment, as in Faig and Jerez (2005), both buyers and sellers are ex ante identical and buyers' private utility shocks occur *after* meetings take place. This approach is more realistic for our application of retail trade and monetary exchange.

[1]Rocheteau and Wright (2009) use a slightly different formulation to calibrate the model in RW. Instead of seller entry, agents can decide whether to be buyers or sellers.

Importantly, meetings are many-on-one or multilateral in our environment, allowing buyers to *choose* sellers within meetings – the key focus of our paper.

Two closely related monetary models that feature buyer utility shocks and private information are Faig and Jerez (2006) and Dong and Jiang (2014). The key novelty of our paper is the fact that consumer choice implies the distribution of chosen goods is *endogenous* and depends on the equilibrium seller-buyer ratio, but this feature in absent from both papers since meetings are bilateral.[2] Faig and Jerez (2006) incorporates money into the environment in Faig and Jerez (2005), which examines the role of private information in a competitive search model of retail trade. Equilibrium features both overconsumption (for lower utility shocks) and underconsumption (for higher utility shocks), but the Friedman rule restores efficiency. Since Faig and Jerez (2006) do not incorporate an individual rationality constraint for buyers, their results are similar to imposing that equilibrium is full-trade in our environment (i.e. all meetings result in trade) and there is no consumer choice.[3] In Dong and Jiang (2014), the trading mechanism is price posting with undirected search.[4] In equilibrium, there is underconsumption of all goods and the Friedman rule does not deliver the efficient allocation. By contrast, in our environment, there are various possibilities for ranges of overconsumption and underconsumption, and the Friedman rule can sometimes deliver efficiency if the equilibrium is full trade.

While Faig and Jerez (2006) and Dong and Jiang (2014) highlight the effect of inflation on non-linear pricing, we focus on the effect of consumer choice on efficiency and the welfare cost of inflation. In that sense, our paper is more similar in spirit to the monetary model of endogenous product variety found in Dong (2010). Dong (2010) considers the effect of product variety on the cost of inflation in an environment with complete information where firms invest to expand product variety. Greater product variety increases welfare not by increasing average quality or utility from consumption, but by increasing the *probability* of trade. Both the equilibrium quantity traded and the measure of product varieties are less than socially optimal. In contrast to our results, the Friedman rule restores efficiency, and the effect of endogenous product variety on the cost of inflation is negligible with competitive search.

[2] Also, both Faig and Jerez (2006) and Dong and Jiang (2014) focus on a uniform distribution of shocks, while our results hold for general exogenous distributions of utility shocks.

[3] In their Appendix, Faig and Jerez (2006) mention how the equilibrium changes if an IR constraint is imposed, but they do not incorporate this possibility into their main results.

[4] In Dong and Jiang (2014), competitive search is discussed in an Appendix.

3 Model

Time is discrete and continues forever. Each period $t \in \{0, 1, 2, ...\}$ is divided into two subperiods, as in Lagos and Wright (2005). During the day, there is a frictionless, centralized market and at night there is a frictional, decentralized market. As in RW, there is a continuum of agents who are divided into two types: *buyers* and *sellers*. Buyers are ex ante identical and sellers are ex ante identical. The sets of buyers and sellers are denoted by B and S respectively. While all agents both produce and consume during the day, buyers and sellers are different at night: buyers wish to consume (but cannot produce) and sellers do not wish to consume (but can produce).

There is a fixed measure of buyers and we normalize $|B| = 1$. All buyers participate in the night market at zero cost, but there is an entry decision by sellers. Only a subset $\bar{S}_t \subseteq S$ of sellers of measure n_t enter the night market. Sellers may or may not choose to enter the night market at cost $k > 0$ and thus n_t is endogenous.[5] Since $|B| = 1$, the measure of sellers who enter, n_t, is also the seller-buyer ratio.

Money is perfectly divisible. The aggregate money supply at date t is $M_t \in \mathbb{R}_+$, which grows at a constant rate γ, i.e. $M_{t+1} = \gamma M_t$. Money is either injected into the economy ($\gamma > 1$) or withdrawn ($\gamma < 1$) by lump sum transfers during the day. We assume these transfers are to buyers only, and we restrict attention to policies where $\gamma \geq \beta$, where β is the discount factor. When $\gamma = \beta$ (the Friedman rule), we only consider equilibria obtained by taking the limit as $\gamma \to \beta$ from above.

The day market is identical to RW, but the night market has two features that enable the possibility of consumer choice. First, meetings are not always bilateral; they can be *many-on-one* or *multilateral*. Each buyer can potentially meet many sellers in a single meeting, but he can choose to trade with only one. Second, the goods offered by sellers differ ex post in terms of buyers' (idiosyncratic) utility shocks. These are random seller-specific utility shocks that reflect individual preferences.

The probability that a buyer meets $N \in \{0, 1, 2...\}$ sellers is given by $P_N(n) = \Pr(N_i = N)$ where $P_N(n) \in [0, 1]$ and $\sum_{N=0}^{\infty} P_N(n) = 1$. We assume that $P_N(n)$ is Poisson, i.e. $P_N(n) = \frac{n^N e^{-n}}{N!}$ for all $N \in \{0, 1, 2...\}$. The probability $\alpha_b(n)$ that a buyer has the opportunity to trade equals the probability that the buyer meets at least one seller, i.e. $\alpha_b(n) = 1 - P_N(0)$, and the probability that a seller is chosen is $\alpha_s(n) = \alpha_b(n)/n$. For simplicity, we denote $\alpha_b(n) = 1 - e^{-n}$ by $\alpha(n)$.

[5] We assume the set S is sufficiently large that $n_t \leq |S|$ always.

At night, sellers can produce on demand any quantity $q \in \mathbb{R}_+$ of a divisible good. Sellers who produce quantity q pay a cost of production $c(q)$. When a buyer meets $N \in \{1, 2, ...\}$ sellers, he draws a seller-specific random utility shock a for each seller he meets and then chooses a single seller with whom to trade. Importantly, these utility shocks are not observed by sellers; they are private information for the buyer.

We sometimes refer to the utility shock a as (perceived) *quality*. A buyer who consumes quantity q of a good with quality a receives utility $au(q)$. The random utility or quality shocks a are drawn from a bounded, continuous distribution with cdf G that is known to all agents. We assume throughout the paper that this distribution is not degenerate and we make the following mild assumptions.

Assumption 1. *The distribution of utility shocks has cdf G that is twice differentiable, pdf $g = G' > 0$, and support $[a_0, \bar{a}]$ where $a_0, \bar{a} \in \mathbb{R}_+$. The distribution has a weakly increasing hazard rate, $\lambda'(a) \geq 0$ where $\lambda(a) \equiv \frac{g(a)}{1-G(a)}$, and $G''(a) \leq 0$.*

The instantaneous utility of a buyer who meets a seller at night at date t is

$$U_t^b = \nu(x_t) - y_t + \beta E_{\tilde{G}_t}(au(q_{a,t})), \tag{1}$$

where x_t is the quantity consumed and y_t is the quantity produced during the day, $q_{a,t}$ is the quantity consumed at night, a is the *quality* of the good consumed, and $\tilde{G}_t$ is the distribution of *chosen goods* at time t. We assume that $u(0) = 0$, $u'(0) = \infty$, $u'(q) > 0$, and $u''(q) < 0$ for all q. We also assume $\nu'(x) > 0$ and $\nu''(x) < 0$ for all x, and there exists x^* such that $\nu'(x^*) = 1$. For now, we normalize $\nu(x^*) - x^* = 0$.[6]

The instantaneous utility of a seller who is chosen by a buyer at night at date t is

$$U_t^s = \nu(x_t) - y_t - \beta E_{\tilde{G}_t}(c(q_{a,t})), \tag{2}$$

where x_t is the quantity consumed and y_t is the quantity produced during the day, and $q_{a,t}$ is the quantity produced at night of a good of quality a. We assume that $c(0) = 0$, $c'(q) > 0$ and $c''(q) \geq 0$ for all $q > 0$.

In the day market, the price of goods is normalized to 1 for all t and the relative price of money is denoted by ϕ_t. In the night market, prices are determined in competitive search equilibrium, which we discuss in Section 6.

[6]Later, when we calibrate the model in Section 8, we will reverse this normalization.

The real value of a quantity of money m_t held by an agent at date t is defined as $z_t \equiv \phi_t m_t$ and the aggregate real money supply is $Z_t \equiv \phi_t M_t$. We will focus on steady-state equilibria where all of the aggregate real variables are constant. Since $M_{t+1}/M_t = \gamma$, this implies that in steady state $\phi_{t+1}/\phi_t = 1/\gamma$.

4 Distribution of chosen goods

While the distribution of the quality of *available goods* G is exogenous, the distribution of the quality of *chosen goods* $\tilde{G}$ is endogenous and depends on both the equilibrium seller-buyer ratio n and the choices made by buyers in equilibrium.

We will later prove that, in any equilibrium, buyers always choose the highest quality seller they meet. Therefore, the endogenous distribution of chosen goods equals the distribution across buyers of the highest quality a among the sellers a buyer meets, conditional on meeting at least one seller. In anticipation of this result, we refer to the latter distribution as the *distribution of chosen goods,* $\tilde{G}$.

The *average quality of a chosen good* is defined by $\tilde{a}_G(n) \equiv E_{\tilde{G}}(a)$, i.e.

$$\tilde{a}_G(n) = \int_{a_0}^{\bar{a}} a \; d\tilde{G}(a;n). \tag{3}$$

For simplicity, we often drop the subscript G and denote $\tilde{a}_G(n)$ simply by $\tilde{a}(n)$, although of course it depends on the underlying distribution G.

Lemma 1 presents an expression for the distribution of chosen goods $\tilde{G}$ and states that, if $n > 0$, the distribution $\tilde{G}$ first-order stochastically dominates the distribution of available goods G. The average quality of a *chosen* good $\tilde{a}(n)$ is thus strictly greater than the average quality of an *available* good.

Lemma 1. *Suppose that the seller-buyer ratio $n > 0$.*

1. *The distribution of chosen goods is given by*

$$\tilde{G}(a;n) = \frac{e^{-n(1-G(a))} - e^{-n}}{1 - e^{-n}}. \tag{4}$$

2. *The distribution of chosen goods $\tilde{G}$ first-order stochastically dominates the distribution of available goods G and $\tilde{a}(n) > E_G(a)$.*

3. In the limit as $n \to 0$, we have $\tilde{G}(a;n) \to G(a)$ and $\tilde{a}(n) \to E_G(a)$.

4. In the limit as $n \to \infty$, we have $\tilde{G}(a;n) \to 0$ for all $a \in [a_0, \bar{a})$ and $\tilde{a}(n) \to \bar{a}$.

Lemma 2 presents a general result that will prove useful. One implication of this result is that the average quality of a chosen good $\tilde{a}(n)$ is strictly increasing in n, i.e. $\tilde{a}'(n) > 0$. Intuitively, the average quality of a chosen good is increasing in the seller-buyer ratio because more sellers per buyer means greater choice for buyers and a higher expected value for the quality of their chosen good.

Lemma 2. *Suppose that the seller-buyer ratio $n > 0$. Let $f(a) : [a_0, \bar{a}] \to \mathbb{R}_+$ and $f' > 0$. The expected value of $f(a)$ with respect to the distribution of chosen goods, $\tilde{f}(n) \equiv \int_{a_0}^{\bar{a}} f(a) d\tilde{G}(a;n)$, is strictly increasing in n, i.e. $\tilde{f}'(n) > 0$.*

Figure 1 compares the density $\tilde{g}(a;n)$ of the distribution of *chosen goods* and the density of the distribution of *available goods* $g(a)$, which is uniform on $[0,1]$ in this example. The density $\tilde{g}(a;n)$ is lower than the density $g(a)$ for low values of a and higher for high values of a. This is because consumer choice shifts the distribution towards the higher quality goods that are actually chosen by buyers.

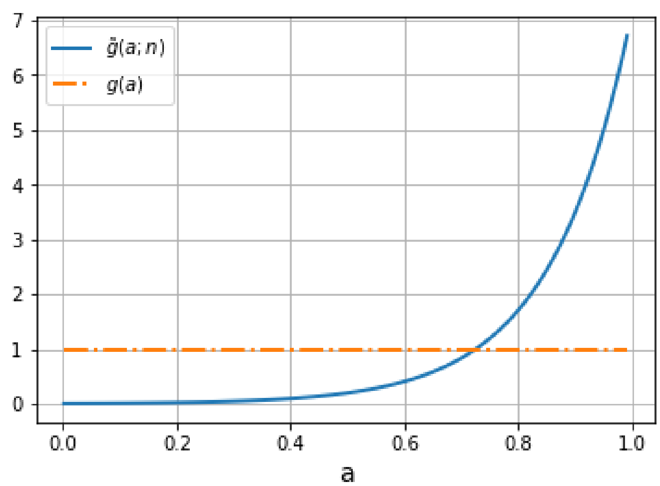

Figure 1: Example of density of the distribution of chosen goods versus available goods

We will later see that, in equilibrium, it may be the case that not all chosen goods are traded. The distribution of *chosen goods* may therefore be different than the distribution of *traded goods*, which we discuss later.

5 Planner's problem

Before we consider competitive search equilibrium, we solve the planner's problem. The cost of seller entry is $k > 0$ and the planner chooses a seller-buyer ratio n^* and a function $q^* : [a_0, \bar{a}] \to \mathbb{R}_+$ to maximize the total surplus created minus the costs of seller entry. That is, the planner solves the following problem:

$$\max_{n \in \mathbb{R}_+, \{q_a\}_{a \in [a_0, \bar{a}]}} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} [au(q_a) - c(q_a)] \, d\tilde{G}(a; n) - nk \right\}. \tag{5}$$

The planner must take into account the fact buyers' expected utility from consumption in the night market depends not only on the meeting probability and the quantity of goods traded, but also on the expected quality of the good purchased. Since the planner is subject to the same search frictions as the decentralized market, we say that the planner's solution achieves the constrained efficient allocation. For brevity, we often refer to this simply as the *efficient allocation*.

Define $s_a \equiv au(q_a) - c(q_a)$, the match surplus for a good of quality a. Let q_a^* denote the efficient quantity of good a and define $s_a^* \equiv au(q_a^*) - c(q_a^*)$. We assume that $s_0^* \geq 0$ where $s_0^* \equiv a_0 u(q_0) - c(q_0)$, so there is (weakly) positive trading surplus for all goods. Let the *expected match surplus* for a chosen good be defined by

$$\tilde{s}(n; \{q_a\}_{a \in [a_0, \bar{a}]}) \equiv \int_{a_0}^{\bar{a}} [au(q_a) - c(q_a)] d\tilde{G}(a; n). \tag{6}$$

For simplicity of notation, throughout the paper, we sometimes suppress the dependence of the expected match surplus $\tilde{s}(n; \{q_a\}_{a \in [a_0, \bar{a}]})$ on the function $q : [a_0, \bar{a}] \to \mathbb{R}_+$ and let $\tilde{s}(n)$ denote $\tilde{s}(n; \{q_a\}_{a \in [a_0, \bar{a}]})$ and $\tilde{s}'(n)$ denote $\partial \tilde{s}(n) / \partial n$.

The following assumption ensures the existence of a social optimum where $n^* > 0$. Intuitively, this condition says that the expected match surplus in the limit as $n \to 0$, i.e. $\lim_{n \to 0} \tilde{s}(n)$, must be greater than k. Since $\tilde{G} \to G$ as $n \to 0$ by Lemma 1, we have $\lim_{n \to 0} \tilde{s}(n) = E_G[au(q_a^*) - c(q_a^*)]$. It follows from our assumptions that, for all $a \in [a_0, \bar{a}]$, there exists a unique $q_a^* \in \mathbb{R}_+$ such that $au'(q_a^*) = c'(q_a^*)$.

Assumption 2. *The cost of entry is not too high:* $E_G[au(q_a^*) - c(q_a^*)] > k$.

Proposition 1 states that there exists a unique social optimum $(n^*, \{q_a^*\}_{a \in [a_0, \bar{a}]})$ with $n^* > 0$ and provides the necessary conditions for an efficient allocation.

Proposition 1. *There exists a unique social optimum* $(n^*, \{q_a^*\}_{a\in[a_0,\bar{a}]})$ *and it satisfies the following necessary conditions:*

1. *For any* $a \in [a_0, \bar{a}]$, *the quantity* $q_a^* > 0$ *solves*

$$au'(q_a^*) = c'(q_a^*). \tag{7}$$

2. *The seller-buyer ratio* $n^* > 0$ *satisfies*

$$\alpha'(n^*)\tilde{s}(n^*; \{q_a^*\}_{a\in[a_0,\bar{a}]}) + \alpha(n^*)\tilde{s}'(n^*; \{q_a^*\}_{a\in[a_0,\bar{a}]}) = k. \tag{8}$$

3. *The distribution of chosen goods is given by (4).*

Equation (8) in Proposition 1 is a version of the *generalized Hosios condition* identified in Mangin and Julien (2020). This condition is a generalization of the well-known Hosios (1990) condition, which states that sellers' surplus share must equal the elasticity of the matching probability $\alpha(n)$ with respect to sellers. To see this equivalence, observe that (8) is equivalent to

$$\frac{\alpha'(n)n}{\alpha(n)} + \frac{\tilde{s}'(n)n}{\tilde{s}(n)} = \frac{nk}{\alpha(n)\tilde{s}(n)}. \tag{9}$$

Defining $\eta_\alpha(n) = \alpha'(n)n/\alpha(n)$ and $\eta_s(n) = \tilde{s}'(n)n/\tilde{s}(n)$, this is equivalent to

$$\underbrace{\eta_\alpha(n)}_{\text{matching elasticity}} + \underbrace{\eta_s(n)}_{\text{surplus elasticity}} = \underbrace{\frac{nk}{\alpha(n)\tilde{s}(n)}}_{\text{sellers' surplus share}}. \tag{10}$$

The term on the right is the sellers' surplus share when there is free entry of sellers at cost k. The generalized Hosios condition states that efficiency of seller entry requires that sellers' surplus share equals the elasticity of the matching probability $\alpha(n)$ plus the elasticity of the expected match surplus $\tilde{s}(n)$ with respect to n. By contrast, when there is no consumer choice, $\tilde{s}(n)$ no longer depends on the seller-buyer ratio n and condition (10) reduces to the standard Hosios condition.

Since s_a^* is increasing in a, Lemma 2 implies that the expected match surplus $\tilde{s}(n)$ is increasing in the seller-buyer ratio n.[7] Therefore, the surplus elasticity $\eta_s(n)$ is

[7] It is established in the proof of Proposition 1 that both q_a^* and s_a^* are increasing in a.

strictly positive, for any $n > 0$. Intuitively, more sellers per buyer means that buyers have greater choice, which increases both the expected quality of chosen goods and the quantities traded (since q_a^* is increasing in a), thus increasing the size of the expected surplus. Therefore, the externality arising from the effect of seller entry on the expected match surplus is always positive. We call this the *quality externality.*

When the generalized Hosios condition (10) holds, entry of sellers is efficient whenever the quantities traded are efficient. This is because both the search externalities and the quality externality (which arises due to consumer choice) are internalized. When this condition holds, sellers are compensated for the effect of seller entry on both the meeting probability and the average quality of chosen goods.

6 Competitive search equilibrium

Competitive search is an equilibrium concept developed in Moen (1997) and Shimer (1996). The basic idea is that either buyers or sellers, or market makers, can post prices or contracts that specify the terms of trade offered. Search is directed in the sense that buyers and sellers choose which *submarket* to enter, where each submarket corresponds to a particular specification of the terms of trade. Commitment is key: buyers and sellers who enter a submarket *commit* to trade at the terms specified within that submarket. Within each submarket, there are search frictions: buyers and sellers in a submarket are matched according to a meeting technology.

We follow the approach of Rocheteau and Wright (2005). There are agents called market makers who can open submarkets by posting terms of trade or contracts. Market makers take into account the expected relationship between the posted terms of trade or contracts and the seller-buyer ratio n. In our environment, market makers post contracts $\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}$ which specify the quantity of the good q_a and the payment in real dollars d_a *contingent on the buyer's utility shock for the chosen seller.*

Agents can search in any submarket they choose, and within each submarket meetings are many-on-one or multilateral. This gives rise to the possibility of consumer choice. After entering a submarket, but before trade occurs, the buyer meets either no sellers, one seller, or many sellers in a meeting. He then draws a utility or "quality" shock for each seller and chooses one seller with whom to trade.

Sellers do not observe buyers' utility shocks, but buyers may choose to reveal their private information to their *chosen seller* through their choice of contract (q_a, d_a)

offered by the seller. By the revelation principle, it is without loss of generality to focus on incentive-compatible direct mechanisms $\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}$ that induce buyers to truthfully reveal their private information to their chosen sellers.

Within each period, the timing is as follows. At the start of each day, the market makers announce the submarkets $\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}$ that will be open that night, implying an expected n for each submarket. During the day, agents trade in the centralized market and readjust their real balances, and then choose a submarket in which to trade at night, in a manner consistent with expectations. During the night, agents trade goods and money in the decentralized market in the submarket of their choice, where they are bound by the posted contracts $\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}$ in that submarket.

Let Ω denote the set of open submarkets, where each submarket $\omega \in \Omega$ is characterized by $(\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}, n)_\omega$, which lists both the contracts offered $\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}$ and the implied seller-buyer ratio n in that submarket.

Let $W^b(z)$ and $W^s(z_s)$ denote the value functions for buyers and sellers respectively in the day market, and let $V^b(z)$ and $V^s(z_s)$ denote the value functions for buyers and sellers respectively in the night market.

Centralized market. In the centralized market, a buyer with real balance z solves:

$$W^b(z) = \max_{\hat{z},x,y\in\mathbb{R}_+} \{\nu(x) - y + \beta V^b(\hat{z})\}, \tag{11}$$

subject to

$$\hat{z} + x = z + T + y, \tag{12}$$

where T is her real transfer and $\hat{z}$ is the real balances carried forward into that period's decentralized night market. Substituting (12) into (11), we get

$$W^b(z) = z + T + \max_{\hat{z},x\in\mathbb{R}_+} \{\nu(x) - x - \hat{z} + \beta V^b(\hat{z})\}. \tag{13}$$

Thus, the buyer's $\hat{z}$ is independent of z, and $W^b(z) = z + W^b(0)$, which is linear.

Similarly, a seller with real balance z_s in the centralized market solves:

$$W^s(z_s) = \max_{\hat{z},x,y\in\mathbb{R}_+} \left\{\nu(x) - y + \beta \max\left[V^s(\hat{z}), W^s\left(\frac{\hat{z}}{\gamma}\right)\right]\right\}, \tag{14}$$

subject to

$$\hat{z} + x = z_s + y. \tag{15}$$

By substituting (15) into (14) we obtain

$$W^s(z_s) = z_s + \max_{\hat{z},x\in\mathbb{R}_+} \left\{ \nu(x) - x - \hat{z} + \beta \max \left[V^s(\hat{z}), W^s\left(\frac{\hat{z}}{\gamma}\right)\right]\right\}. \tag{16}$$

Thus, the seller's $\hat{z}$ is independent of z_s, and $W^s(z_s) = z_s + W^s(0)$.

Decentralized market. For a seller in the decentralized night market,

$$V^s(z_s) = \max_{\omega\in\ \Omega} \left\{ \frac{\alpha(n)}{n}\left[\int_{a_0}^{\bar{a}} -c(q_a) + W^s\left(\frac{z_s + d_a}{\gamma}\right) d\tilde{G}(a;n)\right] + \left[1 - \frac{\alpha(n)}{n}\right] W^s\left(\frac{z_s}{\gamma}\right)\right\} - k \tag{17}$$

where each submarket $\omega \in \Omega$ is characterized by $(\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}, n)$. A seller chooses ω among the set of open submarkets and he has the opportunity to trade if he is chosen. It is straightforward to verify that the seller's choice of real balances is $\hat{z} = 0$.[8]

For a buyer in the decentralized night market,

$$V^b(z) = \max_{\omega\in\ \Omega} \left\{ \alpha(n)\mathbf{1}_{z\geq d_a} \left[\int_{a_0}^{\bar{a}} au(q_a) + W^b\left(\frac{z - d_a}{\gamma}\right) d\tilde{G}(a;n)\right] + [1 - \alpha(n)\mathbf{1}_{z\geq d_a}] W^b\left(\frac{z}{\gamma}\right)\right\}, \tag{18}$$

where $\mathbf{1}_{z\geq d_a}$ is an indicator function that is equal to one if $z \geq d_a$ and zero otherwise. A buyer chooses ω among the set of open submarkets and she gets the opportunity to trade if she meets at least one seller and has sufficient money z to pay the posted d_a for the good she chooses to purchase. Using $W^b(z) = z + W^b(0)$ we obtain

$$V^b(z) = \max_{\omega\in\ \Omega} \left\{ \alpha(n)\mathbf{1}_{z\geq d_a} \int_{a_0}^{\bar{a}} \left[au(q_a) - \frac{d_a}{\gamma}\right] d\tilde{G}(a;n) + \frac{z}{\gamma} + W^b(0)\right\}. \tag{19}$$

Thus, the buyer's choice of z from (13) is given by

$$\max_{z\in\mathbb{R}_+} \left\{ -z + \beta \max_{\omega\in\ \Omega} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \left[au(q_a) - \frac{d_a}{\gamma}\right] d\tilde{G}(a;n) + \frac{z}{\gamma}\right\}\right\} \tag{20}$$

[8] Using $W^s(z_s) = z_s + W^s(0)$, (17) simplifies to $V^s(z_s) = z_s/\gamma + V^s(0)$. Substituting into (16), the choice of $\hat{z}$ is given by the first order condition $-1 + \beta/\gamma \leq 0$, where $-1 + \beta/\gamma = 0$ if $\hat{z} > 0$. Since we only consider the case $\gamma = \beta$ by taking the limit as $\gamma \to \beta$ from above, the solution is $\hat{z} = 0$.

subject to the liquidity constraint, $d_a \leq z$ for all $a \in [a_0, \bar{a}]$.

Defining $i \equiv \frac{\gamma - \beta}{\beta}$, the nominal interest rate, the above problem is equivalent to

$$\max_{z \in \mathbb{R}_+, \omega \in \Omega} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \left[au(q_a) - \frac{d_a}{\gamma} \right] d\tilde{G}(a; n) - i \frac{z}{\gamma} \right\}, \tag{21}$$

subject to $d_a \leq z$ for all $a \in [a_0, \bar{a}]$ plus the constraint that a submarket with posted contracts $\{(q_a, d_a)\}_{a \in [a_0, \bar{a}]}$ will attract measure n of sellers per buyer, where n satisfies

$$\frac{\alpha(n)}{n} \int_{a_0}^{\bar{a}} \left[-c(q_a) + \frac{d_a}{\gamma} \right] d\tilde{G}(a; n) \leq k \tag{22}$$

and $n \geq 0$ with complementary slackness. Due to the presence of private information, we need to impose some additional constraints on problem (21), which we discuss next.

6.1 Existence, uniqueness, and characterization

We focus on incentive-compatible direct mechanisms that induce buyers to reveal their private information to their chosen sellers. Given this, we need to impose on problem (21) two additional constraints: an incentive compatibility (IC) constraint and an individual rationality (IR) constraint. The IR constraint for buyers is

$$au(q_a) - \frac{d_a}{\gamma} \geq 0 \tag{23}$$

for all $a \in [a_0, \bar{a}]$. This condition states that buyers always receive a (weakly) positive *ex post* trade surplus, which means that some meetings may not result in trade. The IC constraint is given by

$$au(q_a) - \frac{d_a}{\gamma} \geq au(q_{a'}) - \frac{d_{a'}}{\gamma} \tag{24}$$

for all $a, a' \in [a_0, \bar{a}]$. Intuitively, this condition states that a buyer with utility shock a cannot do better by choosing a contract $(q_{a'}, d_{a'})$ instead of (q_a, d_a).

We restrict attention to steady-state monetary equilibria where $z > 0$ and $n > 0$. We will later prove that there is a unique solution to the market makers' problem and thus there is only one active submarket in equilibrium. Anticipating this result, we simply denote equilibrium by $(\{(q_a, d_a)\}_{a \in [a_0, \bar{a}]}, z, n)$ and define it as follows.

Definition 1. *A competitive search equilibrium is a list* $(\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}, z, n)$ *where* $(q_a, d_a) \in \mathbb{R}^2_+$ *for all* $a \in [a_0, \bar{a}]$ *and* $z, n \in \mathbb{R}_+\backslash\{0\}$*, such that* $(\{(q_a, d_a)\}_{a\in[a_0,\bar{a}]}, z, n)$ *maximizes (21) subject to the constraints (22) and* $d_a \leq z$ *for all* $a \in [a_0, \bar{a}]$*, plus the the IR constraint (23) and the IC constraint (24).*

Lemma 3 tells us that there may exist a non-empty range of qualities such that trade does not occur in equilibrium, i.e. $q_a = 0$. When the good chosen by a buyer within a meeting falls within this range, we call such meetings *no-trade meetings.* There may also exist a non-empty range of qualities such that buyers' purchases are constrained by their money holdings, i.e. $d_a = z$. When the good chosen by a buyer within a meeting falls within this range, we call such meetings *liquidity constrained.* In liquidity constrained meetings, buyers spend all of their money holdings; in no-trade meetings, buyers spend none of their money; and in all other meetings buyers spend a proportion of their money holdings.

Lemma 3. *In any equilibrium where* $i > 0$*,*

1. *No-trade range. There exists a unique* $a_b \in [a_0, \bar{a}]$ *such that* $q_a = 0$ *for all* $a \in [a_0, a_b]$*, and* $q_a > 0$ *for all* $a \in (a_b, \bar{a}]$.
2. *Liquidity constrained range. There exists a unique* $a_c \in [a_0, \bar{a}]$ *such that* $d_a = z$ *for all* $a \in [a_c, \bar{a}]$*, and* $d_a < z$ *for all* $a \in [a_0, a_c)$*. For any* $a \in [a_c, \bar{a}]$*,* q_a *is constant, i.e.* $q_a = q_{a_c} > 0$.
3. *We have* $a_0 \leq a_b \leq a_c \leq \bar{a}$.

Before presenting Proposition 2, it will be useful to define $\pi(a; n) \equiv 1 - \tilde{G}(a; n)$, the probability that a chosen good has quality greater than a. Applying Lemma 3, the probability that a meeting results in trade is given by

$$\pi(a_b; n) = 1 - \tilde{G}(a_b; n). \tag{25}$$

We also define $\varepsilon_\pi(a; n) \equiv -a\pi'(a; n)/\pi(a; n)$, the elasticity of $\pi(a; n)$ with respect to a, where $\pi'(a; n) \equiv \frac{\partial \pi(a;n)}{\partial a}$. This elasticity can be calculated as follows:

$$\varepsilon_\pi(a; n) = \frac{a\tilde{g}(a; n)}{1 - \tilde{G}(a; n)}. \tag{26}$$

For simplicity, we assume $a_0 = 0$. We also make the following assumption, which ensures the existence of equilibrium with $n > 0$. Intuitively, this condition says that the expected match surplus in the limit as $n \to 0$, i.e. $\lim_{n\to 0} \tilde{s}(n)$, must be greater than k, otherwise no sellers enter. Since $\tilde{G} \to G$ as $n \to 0$ by Lemma 1, we have $\lim_{n\to 0} \tilde{s}(n) = E_G[au(q_a^0) - c(q_a^0)]$ where $q_a^0 \equiv \lim_{n\to 0} q_a(n)$ is given by Lemma 4.[9]

Assumption 3. *The cost of entry is not too high:* $E_G[au(q_a^0) - c(q_a^0)] > k$.

Lemma 4. *Consider the limit as* $n \to 0$. *Let* $q_a^0 \equiv \lim_{n\to 0} q_a(n)$ *and let* a_b^0 *denote the unique solution to* $\varepsilon_\pi^0(a) = 1$ *where* $\varepsilon_\pi^0(a) \equiv \frac{ag(a)}{1-G(a)}$. *For all* $a \in [a_0, a_b]$, $q_a^0 = 0$ *and, for all* $a \in (a_b, \bar{a}]$, q_a^0 *satisfies*

$$\left(a - \frac{1-G(a)}{g(a)}\right) u'(q_a) = c'(q_a). \tag{27}$$

We can now present our main result, which establishes the existence and uniqueness of equilibrium and provides a characterization. It is worth observing that existence of equilibrium does not require any upper bound on inflation.

Proposition 2. *There exists a unique equilibrium for any* $i > 0$, *and it satisfies:*

1. *For any* $a \in [a_0, a_b]$, $q_a = 0$, *and* $d_a = 0$.
2. *For any* $a \in (a_b, a_c]$, *the quantity* $q_a > 0$ *solves:*

$$(a - \phi(a))u'(q_a) = c'(q_a) \tag{28}$$

where

$$\phi(a) = \left(1 - \frac{1}{\delta}\right)\left(\frac{1-\tilde{G}(a;n)}{\tilde{g}(a;n)}\right) - \left(\frac{1}{\delta}\right)\frac{i}{\alpha(n)\tilde{g}(a;n)} \tag{29}$$

and

$$\delta = \frac{1}{1-\varepsilon_\pi(a_b;n)}\left(1 + \frac{i}{\alpha(n)\pi(a_b;n)}\right). \tag{30}$$

Also, $d_a/\gamma = au(q_a) - \int_{a_0}^{a} u(q_x)dx$.

[9] Assumption 3 is more complicated than Assumption 2 because q_a depends on n in equilibrium, but the planner's solution q_a^* is independent of n.

3. For any $a \in [a_c, \bar{a}]$, $q_a = q_{a_c}$ *and* $d_a = d_{a_c}$.

4. The value of a_c *satisfies*

$$\int_{a_c}^{\bar{a}} (a - a_c)\tilde{g}(a;n)da = -\left(1 - \frac{1}{\delta}\right)\int_{a_c}^{\bar{a}} [\tilde{G}(a;n) - \tilde{G}(a_c;n)]da + \left(\frac{1}{\delta}\right)\frac{i\bar{a}}{\alpha(n)}. \tag{31}$$

5. Real money holdings $z > 0$ *is given by* $z = d_{a_c}$.

6. The seller-buyer ratio $n > 0$ *satisfies*

$$\alpha'(n)\tilde{s}(n; \{q_a\}_{a\in[a_0,\bar{a}]}) + \alpha(n)\tilde{s}'(n; \{q_a\}_{a\in[a_0,\bar{a}]}) = k. \tag{32}$$

Also, n *is strictly decreasing in* k.

7. The zero profit condition is satisfied:

$$\frac{\alpha(n)}{n}\int_{a_0}^{\bar{a}} \left[-c(q_a) + \frac{d_a}{\gamma}\right] d\tilde{G}(a;n) = k. \tag{33}$$

8. The distribution of chosen goods is given by (4).

A version of the generalized Hosios condition holds *endogenously* in our environment featuring competitive search since (8) is equivalent to the equilibrium condition (32). The only difference between the equilibrium condition (32) and the planner's condition (8) is that the quantities q_a traded in equilibrium may be different than the efficient quantities q_a^*. Since the expected match surplus $\tilde{s}(n; \{q_a\}_{a\in[a_0,\bar{a}]})$ depends not only on the seller-buyer ratio n but also on the quantities q_a, seller entry is not necessarily efficient. However, seller entry is efficient *provided that the quantity traded is efficient*, i.e. $q_a = q_a^*$ for all $a \in [a_0, \bar{a}]$.

In equilibrium, the endogenous value of a_b may or may not be equal to a_0. If $a_b = a_0$, we refer to the equilibrium as *full trade* because all meetings result in trade.[10] Alternatively, if $a_b > a_0$, we refer to the equilibrium as *partial trade*. In partial trade equilibria, the distribution of traded goods is different than the distribution of chosen

[10]While $q_0 = 0$ since we assume $a_0 = 0$, the distribution G is assumed to have no mass points and therefore the probability that a_0 is the quality of a chosen good is zero.

goods, $\tilde{G}$. The *distribution of traded goods*, denoted by $\tilde{G}_T$, has support $(a_b, \bar{a}]$ and

$$\tilde{G}_T(a;n) = \frac{\tilde{G}(a;n)}{1-\tilde{G}(a_b;n)}. \tag{34}$$

Corollary 1 characterizes full trade equilibria. The expression for the equilibrium quantities traded is simpler and we can present it in closed form.

Corollary 1. *Any full trade* ($a_b = a_0$) *equilibrium satisfies:*

1. *For* $a = a_0$, $q_a = 0$, *and* $d_a = 0$.
2. *For all* $a \in (a_0, a_c]$, *the quantity* q_a *solves:*

$$(a - \phi(a))u'(q_a) = c'(q_a), \tag{35}$$

where

$$\phi(a) = -\left(\frac{i}{\alpha(n)+i}\right)\left(\frac{\tilde{G}(a;n)}{\tilde{g}(a;n)}\right). \tag{36}$$

Also, $d_a/\gamma = au(q_a) - \int_{a_0}^{a} u(q_x)dx$.

3. *Parts 3-8 from Proposition 2 hold.*

6.1.1 Special case: limit as entry cost $k \to 0$

We know from Proposition 2 that n is strictly decreasing in k. In the limiting case where the entry cost $k \to 0$, the equilibrium seller-buyer ratio $n \to \infty$. Lemma 1 implies that as $n \to \infty$, the distribution of chosen goods $\tilde{G}(a;n)$ converges to a degenerate distribution with support $\{\bar{a}\}$. In this limiting case, there are no liquidity-constrained meetings and the equilibrium is full-trade. In all meetings, the quantity traded is $q_a = q_{\bar{a}}$, where $q_{\bar{a}}$ solves the following:

$$\left[\bar{a} + \lim_{n\to\infty}\left(\frac{i}{\alpha(n)+i}\right)\left(\frac{\tilde{G}(\bar{a};n)}{\tilde{g}(\bar{a};n)}\right)\right]u'(q_{\bar{a}}) = c'(q_{\bar{a}}), \tag{37}$$

Taking the limit as $n \to \infty$, the quantity traded $q_{\bar{a}}$ solves $\bar{a}u'(q_{\bar{a}}) = c'(q_{\bar{a}})$, which is the efficient quantity.[11] Since the seller-buyer ratio n is efficient whenever the quantities traded are efficient, we have both $q_a = q_a^*$ for all chosen goods and $n = n^*$. That is, the efficient quantity of goods is traded and seller entry is efficient *regardless* of whether the Friedman rule ($i \to 0$) is imposed. By contrast, in RW, we have $q = q^*$ and $n = n^*$ *only* at the Friedman rule; we do not obtain efficiency as $k \to 0$.

7 Results

In this section, we present our key analytical results. First, we present some results regarding when there is underconsumption or overconsumption relative to the efficient quantity and under-entry or over-entry relative to the efficient level. Next, we consider whether the Friedman rule restores efficiency. Finally, we prove a result regarding non-linear pricing or quantity discounts.

7.1 Under/over consumption and under/over entry

As in RW, there are two margins for efficiency: the *intensive margin* (related to quantity traded or consumption) and the *extensive margin* (related to seller entry). We now consider both margins. Importantly, in our environment with consumer choice, the extensive margin now has two components because seller entry affects both the meeting probability for buyers *and* the size of the expected surplus.

We say that there is *underconsumption* of any good of quality a whenever the quantity traded in equilibrium is less than the efficient quantity, i.e. $q_a < q_a^*$, and there is *overconsumption* whenever $q_a > q_a^*$. We say that there is *under-entry* of sellers when the equilibrium seller-buyer ratio is less than the efficient ratio, i.e. $n < n^*$, and there is *over-entry* whenever $n > n^*$.

7.1.1 Underconsumption and overconsumption

Consider expression (29), which gives us the equilibrium quantities for the trading range that is not liquidity constrained, $a \in (a_b, a_c]$. Given that the efficient quantity satisfies $au'(q_a^*) = c'(q_a^*)$, it is clear that we have *underconsumption* if $\phi(a) > 0$, *overconsumption* if $\phi(a) < 0$, and *efficient* consumption if $\phi(a) = 0$.

[11] Using expression (148) derived in the Appendix, $\lim_{n\to\infty}\left(\frac{\tilde{G}(\bar{a};n)}{\tilde{g}(\bar{a};n)}\right) = \lim_{n\to\infty}\frac{1-e^{-nG(\bar{a})}}{ng(\bar{a})} = 0$.

To better understand expression (29), we can interpret it as a weighted average of two terms, where the endogenous weights are $1/\delta \in (0,1]$ and $1 - 1/\delta \in [0,1)$.

$$\phi(a) = \left(1 - \frac{1}{\delta}\right) \underbrace{\left(\frac{1 - \tilde{G}(a;n)}{\tilde{g}(a;n)}\right)}_{\text{weakly positive, } \geq 0} + \left(\frac{1}{\delta}\right) \underbrace{\frac{-i}{\alpha(n)\tilde{g}(a;n)}}_{\text{negative, } <0} \tag{38}$$

Whether or not we have equilibrium overconsumption or underconsumption for a good of quality a depends on the relative weights given to these two terms, as well as their values at a. If the positive term dominates, we have underconsumption, while if the negative term dominates we have overconsumption. If the two terms exactly offset each other, we have efficient consumption at quality a.

Proposition 3 describes the three equilibrium outcomes in terms of underconsumption or overconsumption ranges for $i > 0$ (as depicted in Figure 2).

Proposition 3. *Let* $a_u \equiv \max\{a_c, a_d\}$ *where*

$$a_d \equiv a_c - \phi(a_c), \tag{39}$$

and let a_p *denote the unique solution to*

$$\tilde{G}(a_p;n) = 1 + \frac{i}{\alpha(n)(1-\delta)}. \tag{40}$$

In any equilibrium where $i > 0$*, there are three possible outcomes:*

1. *If* $a_p \leq a_c$*, there is underconsumption on* (a_0, a_p)*, overconsumption on* (a_p, a_u)*, and underconsumption on* $(a_u, \bar{a}]$*.*

2. *If* $a_p \geq a_c$*, there is underconsumption on* $(a_0, \bar{a}]$*.*

3. *If* $a_b = a_0$*, there is overconsumption on* (a_0, a_d) *and underconsumption on* $(a_d, \bar{a}]$*.*

Proposition 3 implies that in any full-trade equilibrium where $i > 0$, there exists at least some range of overconsumption. Conversely, any equilibrium with no overconsumption, i.e. where $a_p \geq a_c$, must be partial-trade, i.e. $a_b < a_0$.

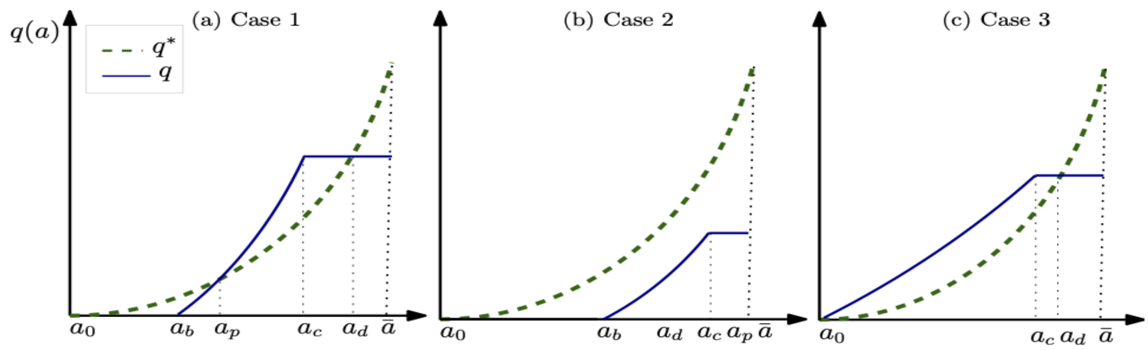

Figure 2: Examples of the three cases of under/over consumption in Proposition 3

7.1.2 Under-entry and over-entry

Given that the generalized Hosios condition holds endogenously under competitive search, we know that the equilibrium seller-buyer ratio n is efficient *provided that the quantities traded* q_a *are efficient*. However, the quantities traded are not efficient whenever $i > 0$ and therefore seller entry is not necessarily efficient.

Proposition 4 states that there can be over-entry, under-entry, or efficient entry of sellers outside the Friedman rule. We can find examples of each possibility.

Proposition 4. *In any equilibrium where* $i > 0$*, there may be either under-entry, over-entry, or efficient entry of sellers.*

While we know that entry must be efficient if the quantity traded is efficient, the converse is not true. There are examples where entry is efficient but the quantities traded are not. When this occurs, the efficiency of entry is really just "coincidental" – both over-entry and under-entry are possible, so entry can also lie in between.

Next, we consider whether imposing the Friedman rule can restore efficiency.

7.2 Does the Friedman rule deliver efficiency?

In RW, there is efficiency along both the intensive and extensive margins when the Friedman rule is imposed. That is, both the quantity traded and the level of entry of sellers are efficient. In our environment, we find that the Friedman rule may or may not deliver efficiency. In fact, there can be inefficiencies along *both* margins.

Corollary 2 characterizes equilibrium at the Friedman rule. The expression for the equilibrium quantities traded is simpler and we can present it in closed form. There are no liquidity-constrained meetings because $a_c = \bar{a}$, but there may still be a range of meetings for which no trade occurs because $a_b > a_0$ is possible.

Recall that $\pi(a; n) \equiv 1 - \tilde{G}(a; n)$, so $\pi(a_b; n)$ is the probability a meeting results in trade, and $\varepsilon_\pi(a; n)$ is the elasticity of $\pi(a; n)$ with respect to a.

Corollary 2. *At the Friedman rule ($i \to 0$), any equilibrium satisfies:*

1. *For any $a \in [a_0, a_b]$, $q_a = 0$, and $d_a = 0$.*

2. *No meetings are liquidity constrained: $a_c = \bar{a}$.*

3. *For all $a \in (a_b, \bar{a}]$, the quantity q_a satisfies*

$$\left(1 - \frac{\varepsilon_\pi(a_b; n)}{\varepsilon_\pi(a; n)}\right) a u'(q_a) = c'(q_a). \tag{41}$$

   Also, $d_a/\gamma = au(q_a) - \int_{a_0}^{a} u(q_x)dx$.

4. *Parts 5-8 from Proposition 2 hold.*

Proposition 5 says the Friedman rule delivers efficiency if and only if the equilibrium is full trade (i.e. $a_b = a_0$). First, it is clear from (41) that the efficient quantity q_a is traded at the Friedman rule if and only if $\varepsilon_\pi(a_b; n) = 0$, which is true if and only if $a_b = a_0 = 0$. Second, the equilibrium condition (32) is equivalent to the planner's condition given the same function q_a, i.e. *given the quantities traded are efficient.*

Proposition 5. *At the Friedman rule, we have efficiency, i.e. $n = n^*$ and $q_a = q_a^*$ for all a, if and only if the equilibrium is full-trade ($a_b = a_0$).*

In general, equilibria will not be full-trade. Thus, the Friedman rule does not generally achieve efficiency; only in the special case where $a_b = a_0$. While possible in theory, this equilibrium outcome rarely occurs. There may be inefficiencies on both margins at the Friedman rule. We consider first the intensive margin (quantities traded) and then the extensive margin (entry of sellers).

7.2.1 Are the efficient quantities traded at the Friedman rule?

Proposition 6 tells us that, in any partial-trade equilibrium where $a_b > a_0$, the Friedman rule results in *underconsumption,* i.e. $q_a < q_a^*$ for all $a \in (a_0, \bar{a})$. Only for two specific qualities of chosen goods, a_0 and $\bar{a}$, are the efficient quantities traded.

Proposition 6. *At the Friedman rule, there is underconsumption for all* $a \in (a_0, \bar{a})$ *if the equilibrium is partial-trade* ($a_b > a_0$).

The reason why the Friedman rule does not yield efficiency along the intensive margin is not only because there is underconsumption in meetings do not result in trade. Even if we consider meetings that *do* result in trade, there is underconsumption. Intuitively, in any partial trade equilibrium, sellers need to compensate for the fact that there is a range of meetings in which no trade occurs. Sellers compensate for the no-trade meetings by charging higher prices over the trading range, which implies that less than the efficient quantity is consumed even within the trading range.

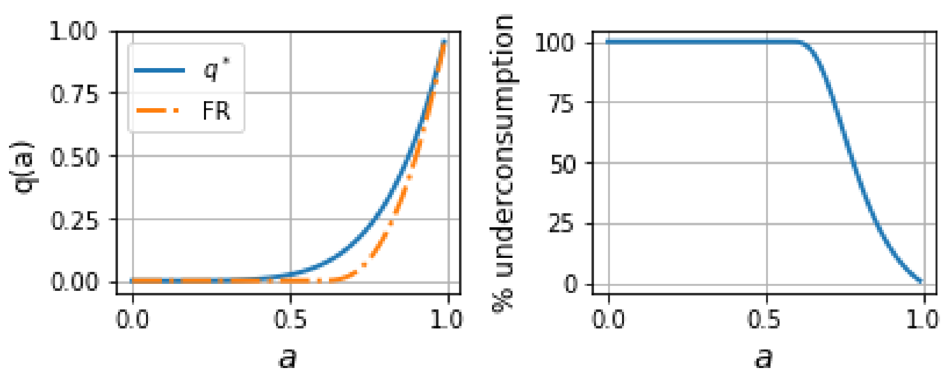

Figure 3: Equilibrium underconsumption at the Friedman rule

Figure 3 presents an example of underconsumption at the Friedman rule ($i \to 0$). The left figure compares the efficient quantity q_a^* and the equilibrium quantity q_a. The right figure depicts the % of underconsumption compared to the efficient quantity.

7.2.2 Is seller entry efficient at the Friedman rule?

Since the generalized Hosios condition holds endogenously under competitive search, we know that the equilibrium seller-buyer ratio n is always efficient *given that the quantities traded* q_a *are efficient*. However, there is generally underconsumption at the Friedman rule (except in the special case where $a_b = a_0$), so n is not necessarily efficient. Therefore, the Friedman rule does not generally deliver efficiency along either the intensive or extensive margin.

Proposition 7. *At the Friedman rule, there can be either under-entry, over-entry, or efficient entry of sellers.*

Proposition 7 says that there may be either under-entry or over-entry of sellers at the Friedman rule. It may also be the case that entry is efficient. In the example depicted in Figure 3, which uses our baseline calibration in Section 8, there is *over-entry* of sellers, i.e. $n > n^*$, at the Friedman rule. We can also find examples for which there is either under-entry or efficient entry.

7.3 Quantity discounts

In the mechanism design literature, it is well-known that when buyers' valuations are private information, sellers offer non-linear price schedules in order to induce buyers with higher valuations to purchase greater quantities (Maskin and Riley, 1984). Consistent with these results, Faig and Jerez (2006) and Dong and Jiang (2014) also find that sellers offer non-linear price schedules in monetary environments with private information. While non-linear pricing schedules due to private information are not the focus of our paper, it is worth confirming that these results extend to our setting.

We say that there is non-linear pricing or *quantity discounts* whenever $p'(q) < 0$, where p is the price per unit and q is the quantity traded. In our environment, we can define the per-unit price as

$$p_a \equiv \frac{d_a/\gamma}{q_a}. \tag{42}$$

Since $q'(a) > 0$ for all $a \in (a_b, a_c)$ or $q \in (0, q_{a_c})$, i.e. on the trading range that is not liquidity-constrained, we can define a differentiable function $p : (0, q_{a_c}) \to \mathbb{R}_+$ that gives the per-unit price $p(q)$ as a function of quantity.[12]

Proposition 8 provides sufficient conditions for quantity discounts, i.e. $p'(q) < 0$. Since we know from Proposition 2 that n is strictly decreasing in k, condition (43) holds provided that k is sufficiently low.

Proposition 8. *Suppose that* $\frac{c'(q)q}{c(q)} \leq 1$ *and* $\frac{u'(q)q}{u(q)} < 1$ *for all* $q \in \mathbb{R}_+$.

1. *There are quantity discounts on any overconsumption range in* (a_b, a_c).

2. *There are quantity discounts at* a, *in any underconsumption range in* (a_b, a_c), *if*

$$g'(a) + ng(a)^2 \geq g(a). \tag{43}$$

[12] Note that for any $a \in [a_0, a_b]$ or $a \in [a_c, \bar{a}]$, $p'(q)$ is undefined since $q'(a) = 0$.

In Section 8, when we calibrate the model, we confirm that quantity discounts arise in equilibrium under our baseline calibration where condition (43) holds for all a. While in theory it is possible there may exist examples where quantity discounts do not occur – at least for some range of underconsumption where condition (43) does not hold – we cannot find any such examples.

8 Quantitative analysis

In this section, we calibrate the model using data on money demand. After calibrating the model, we provide some comparative statics results and then examine the implications of consumer choice for the welfare cost of inflation.

8.1 Calibration

We calibrate the model to match the data from Lucas and Nicolini (2015) on money demand in the U.S. from 1915-2008. Money demand $L(i)$ is defined as $M1/GDP$. Lucas and Nicolini (2015) adjust the measure of $M1$ in order to generate a stable money demand curve. Defining $\tilde{d}(n) \equiv \int_{a_0}^{\bar{a}} \frac{d_a}{\gamma} d\tilde{G}(a;n)$, the average payment for a chosen good, money demand is given by

$$L(i) = \frac{z}{A + \alpha(n)\tilde{d}(n)}. \quad (44)$$

The period length is set to one year. We set $\beta = 1/(1+r)$ to match a real interest rate of $r = 0.03$ as in Bethune et al. (2020). We use the 3-month U.S. T-bill rate as a measure of the nominal interest rate i. The average nominal interest rate i for the period 1915-2008 is $i = 0.0383$ and the average money demand is $L(i) = 0.272$. The elasticity of money demand $L(i)$ with respect to i is -0.16.

We assume that $G(a)$ is uniform on $[0,1]$. We assume $c(q) = q$ and

$$u(q) = \frac{(q+\epsilon)^{1-\sigma} - \epsilon^{1-\sigma}}{1-\sigma} \quad (45)$$

where $\sigma \in (0,1)$ and $\epsilon \approx 0$. We undo the normalization $\nu(x^*) - x^* = 0$. Instead, the CM utility function is $\nu(x) = A \log x$. Since $\nu'(x^*) = 1$, we have $x^* = A$.

We calibrate three parameters (A, σ, k) to match three targets. The steady state level of money demand in the model, $L(i)$ where $i = 0.0383$, is set equal to the average money demand in the data from 1915-2008, and the elasticity of the model's money demand curve is set equal to the elasticity in the data from 1915-2008. Defining the expected buyers' surplus by $\tilde{v}(n) \equiv \int_{a_0}^{\bar{a}} v_a d\tilde{G}(a;n)$ where $v_a \equiv au(q_a) - \frac{d_a}{\gamma}$, *buyers' surplus share* is $\theta(n) \equiv \tilde{v}(n)/\tilde{s}(n)$. We treat $\theta(n)$ as a proxy for buyers' bargaining power and we target $\theta(n) = 0.5$. The welfare cost of inflation is sensitive to this target, so we later vary this target in Section 8.3.5 and show how the results change.

Table 1 reports the calibrated parameters as well as the key endogenous variables (both targets and other variables). The average quality, quantity, payment, and match surplus are expected values with respect to the distribution of chosen goods, $\tilde{G}(a;n)$. For example, we define $\tilde{q}(n) \equiv \int_{a_0}^{\bar{a}} q_a d\tilde{G}(a;n)$.[13]

Calibrated parameters	
DM utility curvature, $1-\sigma$	0.82
CM utility parameter, A	1.75
cost of entry, k	0.008
Calibration targets	
money demand, $L(i)$	0.27
elasticity of money demand	-0.16
buyers' surplus share, $\theta(n)$	0.50
Other variables	
seller/buyer ratio, n	7.16
buyers' meeting prob, $\alpha(n)$	1.00
average quality, $\tilde{a}(n)$	0.86
average quantity, $\tilde{q}(n)$	0.36
average payment, $\tilde{d}(n)$	0.42
average surplus, $\tilde{s}(n)$	0.12
money holdings, z/γ	0.59
% no-trade meetings	0.04
% constrained meetings	0.49

Table 1: Baseline calibration

In our baseline calibration, the equilibrium features underconsumption of goods of *all* qualities (i.e. there is no overconsumption). The equilibrium is *partial trade*. Around 4% of meetings do not result in any trade, while almost 50% of meetings are

[13]The average values across all *traded* goods can be obtained by dividing by $1 - \tilde{G}(a_b;n)$, or one minus the proportion of no-trade meetings.

liquidity constrained. Buyers spend around 70% of their money holdings on average.

We do not target the output share of the decentralized market, but it is worth noting that it is around 17% in our calibration. In the literature, values of the DM output share vary significantly, from less than 10% in Lagos and Wright (2005) to 25% in Bethune et al. (2020) and 42% in Berentsen, Menzio, and Wright (2011). We also do not target the aggregate markup μ, but it is quite low in our calibration: $\mu = 1.03$.[14] This is one limitation of competitive search, which yields low markups compared to environments featuring Nash or Kalai bargaining, where the bargaining parameter can simply be adjusted to match the aggregate markup.

Figure 4 shows the fit of the model to the money demand data from 1915-2008.

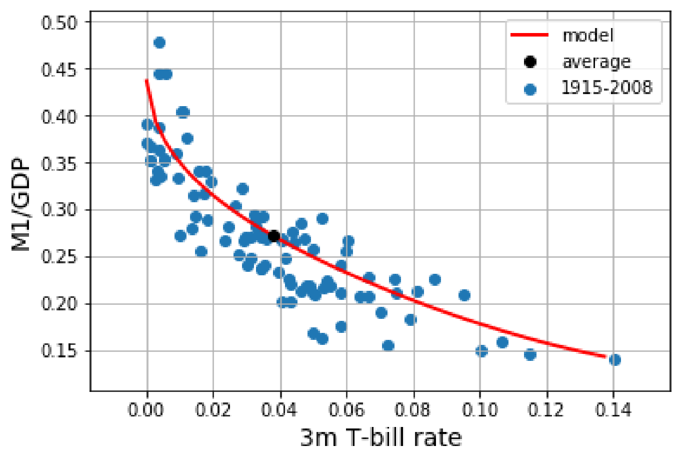

Figure 4: Data versus model predictions for money demand

8.2 Comparative statics

In order to see how changes in the parameters affect the endogenous variables, we now provide some comparative statics results for the baseline calibration with consumer choice. We focus on the effects of changes in the cost of seller entry k and the inflation rate, $\tau \equiv \gamma - 1$.

[14]The DM markup is given by $\mu_{DM} = \tilde{d}(n)/\tilde{q}(n)$ and the CM markup is one. The aggregate markup μ is an average of the two markups that is weighted by output shares.

8.2.1 Comparative statics for inflation rate

Figure 5 presents the comparative statics with respect to the inflation rate τ for the average quality $\tilde{a}(n)$, the average quantity $\tilde{q}(n)$, the average match surplus $\tilde{s}(n)$, and buyers' surplus share, $\theta(n)$. As the inflation rate rises, seller entry declines and the seller-buyer ratio n falls. As a result, the average quality of a chosen good decreases. The average quantity traded and the average match surplus also decline. At the same time, buyers' surplus share decreases due to the fact that higher inflation leads to a lower seller-buyer ratio and therefore a lower effective bargaining power for buyers.

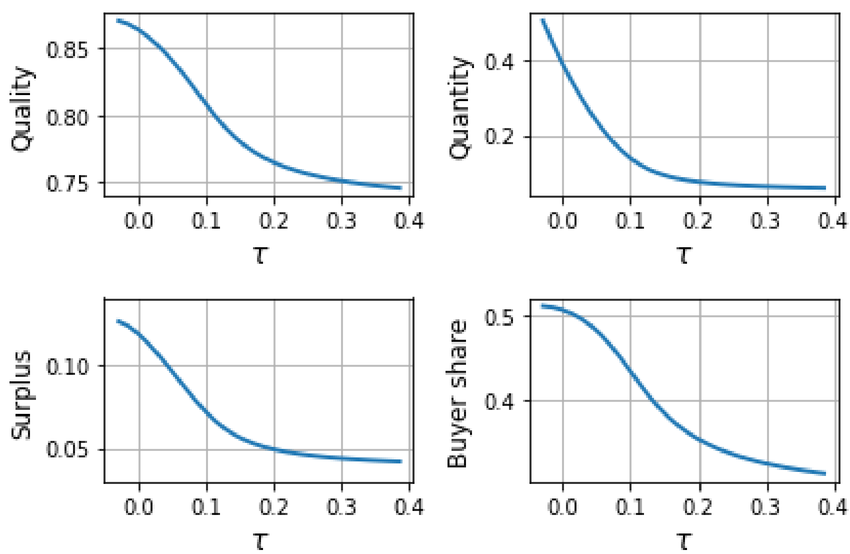

Figure 5: Comparative statics with respect to inflation τ of average quality, average quantity, average surplus, and buyers' surplus share

Figure 6 presents the comparative statics with respect to the inflation rate τ for money holdings z/γ, average payments $\tilde{d}(n)$, the cut-off quality a_c, and the proportion of meetings that are liquidity constrained. As inflation rises, buyers' money holdings and average payments are strictly decreasing. The cut-off quality a_c at which the liquidity constraint binds decreases at first and then increases slightly at very high inflation rates.

Interestingly, the share of meetings that are liquidity-constrained varies *non-monotonically* with the inflation rate τ. This non-monotonicity results from two opposing effects that arise when there is consumer choice. First, as inflation τ increases, money holdings z/γ decrease – a standard effect due to the higher cost of

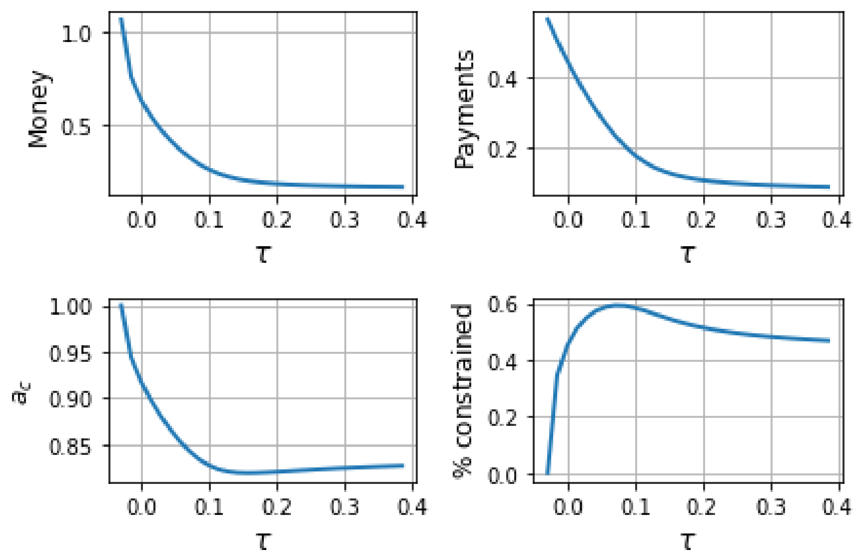

Figure 6: Comparative statics with respect to inflation τ of money holdings, average payments, cut-off quality a_c, and % of liquidity-constrained meetings

holding money. The second, more novel, effect is that as inflation τ increases, the seller-buyer ratio n is lower and therefore the average quality is lower since $\tilde{a}'(n) > 0$. As a result, buyers wish to consume lower quantities and there is less need for liquidity. The first effect dominates for lower inflation rates where the decline in money holdings is steepest, but the second effect dominates for higher inflation rates where money holdings decrease more slowly.

Figure 7 presents comparative statics with respect to inflation τ for the quantity q_a, surplus s_a, and buyers' surplus share θ_a, where $\theta_a \equiv v_a/s_a$, as well as the density of the distribution of chosen goods, $\tilde{g}(a;n)$. There are three cases for inflation τ: low (Friedman rule), baseline ($\tau = 0.008$), and high (10% inflation).

As inflation rises, the seller-buyer ratio n decreases, shifting the density $\tilde{g}(a;n)$ so that lower values of a have higher probability mass and higher values of a have lower probability mass due to the lower degree of consumer choice. As inflation rises, the quantity traded q_a, the surplus s_a, and buyers' surplus share θ_a are all lower for all a. This is largely due to the fact that higher inflation leads to lower entry of sellers.

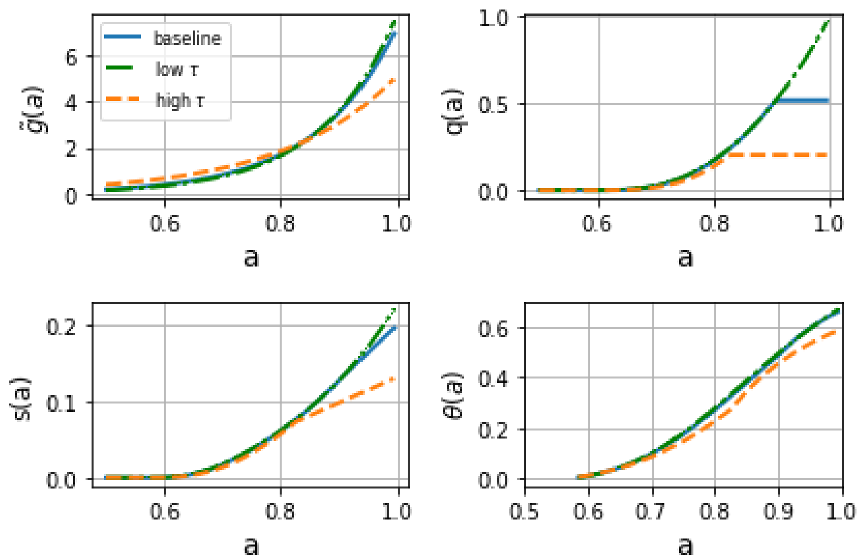

Figure 7: Comparative statics with respect to inflation τ of quantity, surplus, buyers' surplus share, and the density of the distribution of chosen goods, as functions of quality a

8.2.2 Comparative statics for cost of entry

Figure 8 presents the comparative statics with respect to the cost of entry k for the average quality $\tilde{a}(n)$, the average quantity $\tilde{q}(n)$, the average match surplus $\tilde{s}(n)$, and the average buyers' surplus share, $\theta(n)$. As the cost of entry k rises, the average quantity traded falls but also the average quality of chosen goods falls and the average surplus falls, due to the fact that seller entry decreases. Buyers' surplus share is also decreasing with the entry cost k over this range due to the lower seller-buyer ratio leading to lower effective bargaining power for buyers.

Figure 9 presents the comparative statics with respect to the cost of entry k for money holdings z/γ, average payments $\tilde{d}(n)$, the cut-off quality a_c, and the proportion of meetings that are liquidity constrained. As the entry cost k rises, the seller-buyer ratio n declines, average payments decrease due the lower average quality of chosen goods, and the share of constrained meetings falls.

The behavior of money holdings and the cut-off quality a_c is more interesting. Buyers' money holdings and the cut-off a_c vary *non-monotonically* with the entry cost k. Since $z = d_{a_c}$ and d_a is weakly increasing in a, money holdings z/γ are affected by movements in both average payments $\tilde{d}(n)$ and the cut-off a_c. Initially,

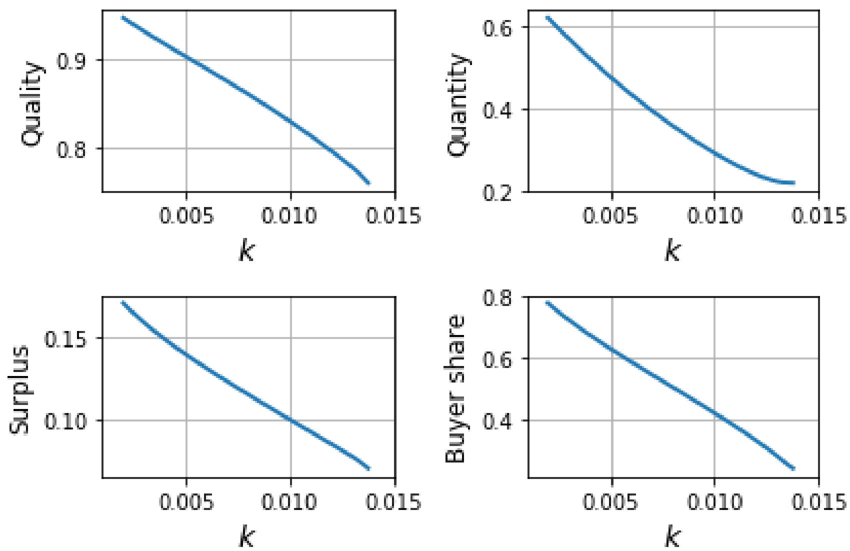

Figure 8: Comparative statics with respect to entry cost k of average quality, average quantity, average surplus, and buyers' surplus share

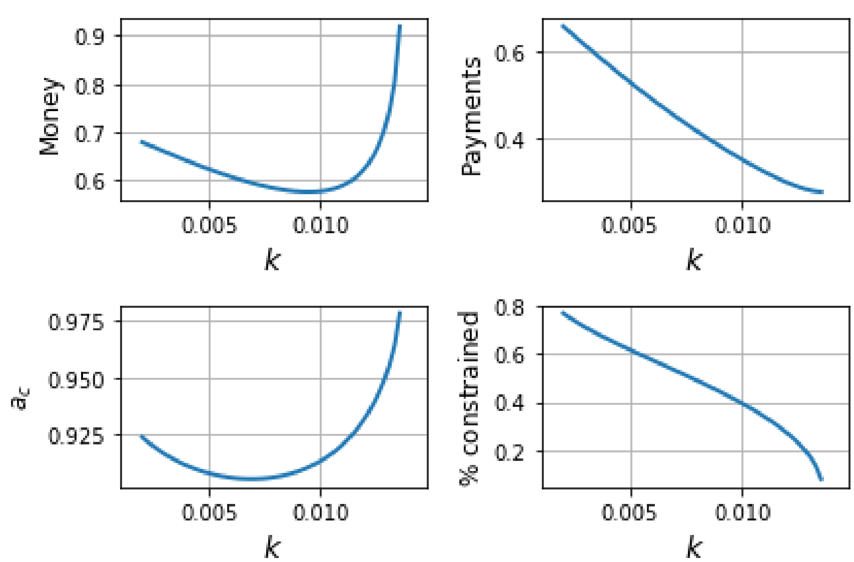

Figure 9: Comparative statics with respect to entry cost k of money holdings, average payments, cut-off quality a_c, and % of liquidity-constrained meetings

as k increases, money holdings fall because average payments $\tilde{d}(n)$ decrease due to the lower average quality of chosen goods. However, when k is even higher and n falls further, fewer meetings are liquidity constrained and a_c increases, leading z/γ to increase as this effect starts to dominate. Eventually, a_c approaches its maximum $\bar{a}$ and z/γ moves closer to its upper bound $\bar{a}u(q_{\bar{a}})$ and then starts to slowly decline (not shown in Figure 9). This non-monotonic behavior of money holdings contrasts markedly with RW, where z/γ is strictly decreasing in k.

Figure 10 presents comparative statics results with respect to the cost of entry k for quantity q_a, surplus s_a, buyers' surplus share θ_a, and the density $\tilde{g}(a;n)$. There are three cases for k: low ($k = 0.004$), baseline ($k = 0.008$), and high ($k = 0.012$).

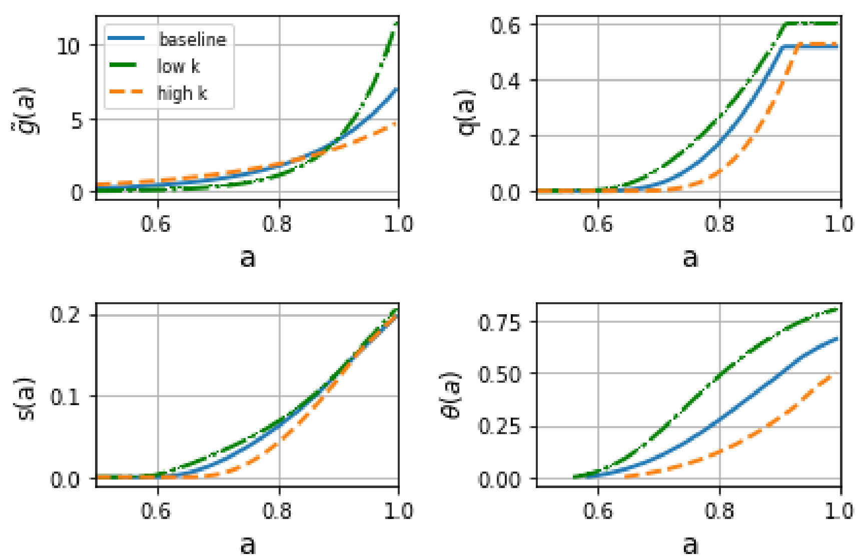

Figure 10: Comparative statics with respect to entry cost k of quantity, surplus, buyers' surplus share, and the density of the distribution of chosen goods, as functions of quality a

As the cost k rises, the seller-buyer ratio n decreases, shifting the density $\tilde{g}(a;n)$ so that higher values of a have lower probability mass due to the lower extent of choice. As k rises, the quantity q_a , surplus s_a, and buyers' surplus share θ_a are generally lower – due to the lower entry of sellers. However, there is some non-monotonicity for values of a above the liquidity cut-off a_c. This is because the cut-off a_c is itself non-monotonic in k, as discussed above.

8.2.3 Comparative statics for quantity discounts

We can consider the effect of both inflation and entry cost on the function $p(q)$. Figure 11 depicts the function $p(q)$ for different values of τ and k (low, baseline, and high). For all values of τ and k depicted in Figure 11, it is clear that $p'(q) < 0$, i.e. there are quantity discounts in equilibrium. Since the conditions in Proposition 8 hold in all cases, this is consistent with that result.

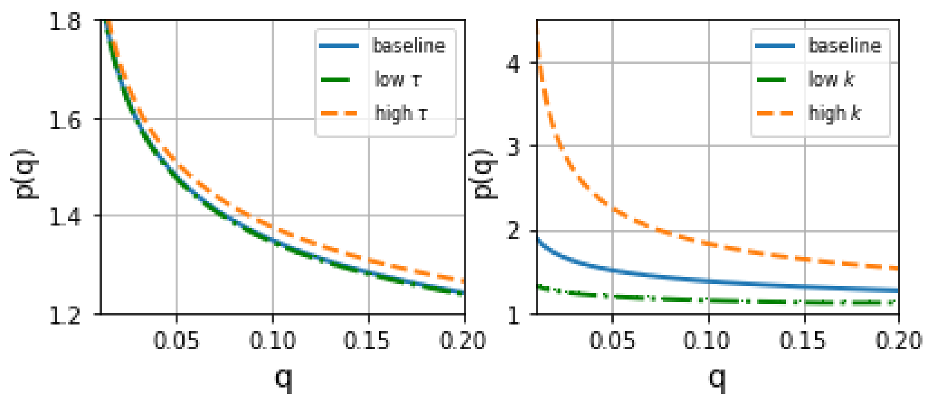

Figure 11: Price per unit as a function of quantity traded

When inflation is higher, the slope of the function $p(q)$ is slightly *flatter*, i.e. the extent of the quantity discounts decreases. This is consistent with results in Faig and Jerez (2006) and Dong and Jiang (2014). Since our model features entry, we can also examine the effect of entry cost. As Figure 11 shows, changes in inflation have only a slight effect on the extent of quantity discounts, but changes in the cost of entry have a significant effect. When the cost of entry is higher (and n lower), the slope of the function $p(q)$ is *steeper*, i.e. the extent of quantity discounts increases. When the cost of entry is lower (and n higher), the function $p(q)$ becomes relatively flat.

8.3 Welfare estimates

In this section, we calculate the welfare cost of inflation. In addition, since the Friedman rule does not deliver efficiency in our environment, we also calculate the welfare loss at the Friedman rule (compared to the efficient allocation).

8.3.1 Measure of welfare cost

When we calculate both the welfare cost of inflation and the welfare loss at the Friedman rule, total welfare in economy E is defined as:

$$W(E) = \alpha(n) \int_{a_0}^{\bar{a}} [au(q_a) - c(q_a)] d\tilde{G}(a; n) - nk + \nu(x^*) - x^*. \tag{46}$$

Since consumers' utility depends on both quality and quantity in our environment, in order to calculate the welfare cost as a consumption sacrifice in terms of quantity alone we first convert to a welfare-equivalent "representative" economy in which the quantity of goods traded is constant and quality is normalized to one. That is, we find the quantity q such that

$$W(E) = \alpha(n)[u(q) - c(q)] - nk + \nu(x^*) - x^*. \tag{47}$$

If total consumption is reduced by a factor of $\Delta \in [0, 1]$, then welfare is given by

$$W(E, \Delta) = \alpha(n)[u(\Delta q) - c(q)] - nk + \nu(\Delta x^*) - x^*. \tag{48}$$

We measure the welfare cost of moving from economy E to economy E' by the share of total consumption that consumers are willing to give up in order to go from economy E' to E. That is, the cost is $1 - \Delta$ where $\Delta \in [0, 1]$ satisfies $W(E, \Delta) = W(E')$.

8.3.2 Effect of consumer choice

In order to examine the effect of consumer choice on the welfare cost of inflation, we generalize our model to one in which the degree of choice may vary between zero and one. Our model with consumer choice in all meetings is a special case.

Let $\pi \in (0, 1]$ be an exogenous parameter that represents the *degree of choice*, i.e. the proportion of meetings that feature consumer choice. In meetings with choice, buyers are informed about their utility shocks *prior* to selecting sellers. We can interpret meetings without choice as meetings in which buyers are informed about their utility shocks only *after* selecting sellers (but before trade).

In meetings with choice, buyers choose to trade with the highest quality seller they meet, as before. However, in meetings without choice, sellers are selected at random. If $\pi = 1$, all meetings feature choice and we recover our main model. As

$\pi \to 0$, all buyers are randomly allocated to sellers, which is effectively equivalent to an environment where all meetings are one-on-one or bilateral.

For any $\pi \in (0, 1]$, the equilibrium can be characterized by exactly the same equations as the model with full choice, found in Proposition 2, except the general distribution of chosen goods $\tilde{G}_\pi(a; n)$ is given by

$$\tilde{G}_\pi(a; n) = \pi\tilde{G}(a; n) + (1 - \pi)G(a). \tag{49}$$

That is, $\tilde{G}_\pi(a; n)$ is a weighted average of the cdf that arises when consumers choose the best seller they meet and the cdf when sellers are chosen at random.

To isolate the effect of consumer choice on the welfare cost of inflation, we take the limit as $\pi \to 0$ and recalibrate the model using the same targets. Table 2 reports the calibrated parameters as well as the targets and other endogenous variables.

Calibrated parameters	
DM utility curvature, $1 - \sigma$	0.63
CM utility parameter, A	2.13
cost of entry, k	0.039
Calibration targets	
money demand, $L(i)$	0.27
elasticity of money demand	-0.16
buyers' surplus share, $\theta(n)$	0.50
Other variables	
seller/buyer ratio, n	1.26
buyers' meeting prob, $\alpha(n)$	0.72
average quality, $\tilde{a}(n)$	0.50
average quantity, $\tilde{q}(n)$	0.14
average payment, $\tilde{d}(n)$	0.21
average surplus, $\tilde{s}(n)$	0.14
money holdings, z/γ	0.62
% no-trade meetings	0.32
% constrained trades	0.18

Table 2: Baseline calibration: no consumer choice

8.3.3 Welfare cost of inflation

To calculate the welfare cost of 10% inflation, we find the value $\Delta_0 \in [0, 1]$ such that $W(\gamma = 1, \Delta_0)$ is equal to $W(\gamma = 1.1, \Delta = 1)$. This is the percentage of total consumption that consumers are willing to give up in order to go from 10% inflation to 0% inflation. We also find the value $\Delta_F \in [0,1]$ such that $W(\gamma = \beta, \Delta_F)$ is equal to $W(\gamma = 1.1, \Delta = 1)$. This is the percentage of total consumption that consumers are willing to give up in order to go from 10% inflation to the Friedman rule.

Without consumer choice (i.e. as $\pi \to 0$), we estimate that the welfare cost of increasing inflation from 0% to 10% is 0.60% of consumption. With consumer choice ($\pi = 1$), the cost of increasing inflation from 0% to 10% is more than twice as high at 1.45% of consumption. Without choice, the cost of moving from the Friedman rule to 10% inflation is 0.78% of consumption, but with choice the cost is again more than twice as high at 1.65% of consumption.

Figure 12 shows how the cost of increasing inflation from 0% to 10% varies with the degree of choice, π. The figure on the left shows the comparative static effect of varying π at the baseline calibration with choice. The figure on the right shows how the cost of inflation varies when we recalibrate the model for different values of π. In both cases, the welfare cost of inflation is strictly increasing in the degree of choice.

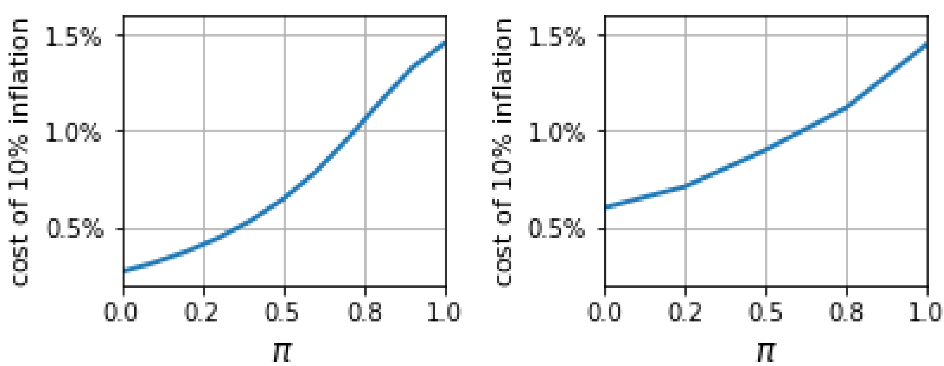

Figure 12: Welfare cost of 10% inflation as a function of degree of choice, π

Why is the cost of inflation higher with consumer choice? We now examine, in greater detail, why there is a difference in the welfare cost of inflation when we have consumer choice. Tables 3 demonstrates how the endogenous variables change when there is consumer choice and the inflation rate changes from the Friedman rule to 0% inflation or 10% inflation. Table 4 presents the corresponding values when there is no consumer choice. In all cases, the model parameters are the same as the baseline calibrations described in Tables 1 and 2, except that the inflation rate varies.

In all cases, the equilibrium is partial trade and there is underconsumption for all $a \in (a_0, \bar{a}]$. We also include the efficient outcomes for comparison.

	Efficient	Friedman rule	0% inflation	10% inflation
buyers' surplus share, $\theta(n)$	-	0.51	0.51	0.43
seller/buyer ratio, n	7.33	7.70	7.31	5.04
buyers' meeting prob, $\alpha(n)$	1.00	1.00	1.00	0.99
average quality, $\tilde{a}(n)$	0.86	0.87	0.86	0.81
average quantity, $\tilde{q}(n)$	0.56	0.50	0.39	0.14
average payment, $\tilde{d}(n)$	-	0.56	0.45	0.18
average surplus, $\tilde{s}(n)$	0.13	0.13	0.12	0.07
money holdings, z/γ	-	1.07	0.64	0.26
% no-trade meetings	0.00	0.04	0.04	0.10
% constrained trades	0.00	0.00	0.45	0.59

Table 3: Equilibrium outcomes for different inflation rates (with choice)

	Efficient	Friedman rule	0% inflation	10% inflation
buyers' surplus share, $\theta(n)$	-	0.52	0.50	0.46
seller/buyer ratio, n	1.40	1.31	1.27	1.12
buyers' meeting prob, $\alpha(n)$	0.75	0.73	0.72	0.68
average quality, $\tilde{a}(n)$	0.50	0.50	0.50	0.50
average quantity, $\tilde{q}(n)$	0.27	0.19	0.15	0.09
average payment, $\tilde{d}(n)$	-	0.26	0.22	0.16
average surplus, $\tilde{s}(n)$	0.16	0.15	0.14	0.12
money holdings, z/γ	-	1.18	0.66	0.41
% no-trade meetings	0.00	0.30	0.32	0.37
% constrained trades	0.00	0.00	0.17	0.27

Table 4: Equilibrium outcomes for different inflation rates (without choice)

Consumer choice leads to a significantly higher welfare cost of inflation due to two main factors: *(i)* the greater effect of inflation on seller entry; and *(ii)* the stronger implications for welfare of any given decline in the seller-buyer ratio. To see this, consider the effect of a change in inflation from 0% to 10% in Tables 3 and 4.

When inflation rises from 0% to 10%, the measure n of sellers who enter falls by 31% with choice compared to a much smaller drop of 12% without choice. With consumer choice, this change in the seller-buyer ratio n leads to a slight fall in the meeting probability for buyers and a larger change in the average quality of a chosen good, which declines by around 6%. Without choice, the meeting probability for buyers falls by around 6%, but average quality is constant. At the same time, the average quantity falls dramatically when there is consumer choice, decreasing by 64%, but less dramatically without choice, falling by around 37%.

Overall, due to changes in both quantity and quality, the average match surplus decreases by 39% when there is consumer choice, but only by 14% without choice. This is the main driver of the difference in welfare costs. To see this, consider the net surplus in the decentralized market, $W_{DM}(\gamma) = \alpha(n)\tilde{s}(n) - nk$. When there is consumer choice, there is a much greater change in the surplus $\alpha(n)\tilde{s}(n)$, which falls by 40% with choice but only 19% without choice. There is also a greater change in total entry costs nk, but the effect on the surplus $\alpha(n)\tilde{s}(n)$ dominates.

8.3.4 Welfare loss at the Friedman rule

The Friedman rule does not deliver efficiency in either of the calibrated versions of our model (with and without choice). However, welfare is maximized at the Friedman rule in both versions.[15] To enable comparison of the equilibrium at the Friedman rule with the efficient allocation, Tables 3 and 4 also include the efficient values of the relevant endogenous variables (i.e. the planner's solutions).

Both when there is consumer choice and when there is no choice, we have under-consumption at the Friedman rule, despite the fact there are no meetings that are liquidity-constrained since $a_c = 1$. When there is no choice, there is also under-entry of sellers. However, when there is choice, there is *over-entry* of sellers, i.e. the efficient seller-buyer ratio is lower than the equilibrium seller-buyer ratio at the Friedman rule. That is, the Friedman rule delivers both *underconsumption* and *over-entry* of sellers relative to the efficient outcome. The fact that there is over-entry of sellers is consistent with Proposition 6. Without consumer choice, over-entry is not possible: there is always under-entry of sellers at the Friedman rule.

Since the Friedman rule does not deliver efficiency in our environment, it is useful to calculate the welfare loss at the Friedman rule (compared to the efficient outcome). To do so, we find the value $\Delta_E \in [0, 1]$ such that total welfare at the efficient outcome, $W(E, \Delta_E)$, is equal to total welfare at the Friedman rule, $W(\gamma = \beta, \Delta = 1)$. This is the percentage of total consumption that consumers are willing to give up in order to go from the Friedman rule to the efficient outcome. While the planner's solution

[15] While we have no proof that welfare is always maximized at the Friedman rule (even though it does not yield the planner's allocation), numerical examples suggest that this is the case.

cannot be decentralized in our environment, this is still a useful way of measuring the extent of the inefficiency at the Friedman rule.

Without consumer choice, we estimate that the cost of moving from the efficient outcome to the Friedman rule is 0.47% of consumption. With consumer choice, we estimate that the cost of moving from the efficient outcome to the Friedman rule is just 0.16% of consumption. Thus, the presence of consumer choice can alleviate the extent of the welfare loss at the Friedman rule.

Why is the welfare loss at Friedman rule lower with choice? The fact that there is over-entry of sellers when there is consumer choice is what drives the difference in welfare losses at the Friedman rule. When we shift from the efficient allocation to the Friedman rule equilibrium, the measure n of sellers *rises* by around 5% with consumer choice, but decreases by around 7% without choice. With consumer choice, this increase in the seller-buyer ratio leads to a slight *increase* in the average quality of a chosen good and a smaller drop in the average quantity traded, which falls by only 10% with choice compared to 30% without choice.

Overall, due to changes in both quantity and quality, the average match surplus decreases by around 9% without choice but is relatively constant with choice, falling by only 0.2%. This accounts for the difference in welfare loss at the Friedman rule. To see this, consider the DM net surplus, $W_{DM}(\gamma) = \alpha(n)\tilde{s}(n) - nk$. Without choice, there is a much greater change in the surplus $\alpha(n)\tilde{s}(n)$, which falls by around 12% but hardly changes with choice. Total entry costs nk increase slightly with choice and decrease without choice, but the effect on the surplus $\alpha(n)\tilde{s}(n)$ dominates. In this way, the possibility of seller over-entry in the presence of consumer choice can alleviate the extent of the welfare loss at the Friedman rule.

8.3.5 Varying buyers' surplus share

Since we focus on competitive search, our estimates of the welfare cost of inflation can be viewed as lower bounds when compared to environments featuring bargaining. In such environments, the cost of inflation is sensitive to changes in the bargaining parameter. In Lagos and Wright (2005), the cost of inflation *decreases* as buyers' bargaining parameter θ increases because the severity of the hold-up problem decreases as $\theta \to 1$. In environments such as ours that feature competitive search, there is no hold-up problem. Buyers' surplus share is endogenous and depends crucially on the equilibrium seller-buyer ratio. Therefore, the expected effect on the cost of inflation of varying the target buyers' surplus share $\theta(n)$ is not obvious.

Table 5 reports the welfare cost of inflation when we vary the target value of buyers' surplus share and recalibrate the model to match the same targets.

	$\theta(n) = 0.4$	$\theta(n) = 0.5$	$\theta(n) = 0.6$
Choice			
Welfare cost of 0% to 10% inflation	0.86%	1.45%	1.60%
Welfare cost of FR to 10% inflation	1.08%	1.65%	1.78%
No choice			
Welfare cost of 0% to 10% inflation	0.28%	0.60%	0.82%
Welfare cost of FR to 10% inflation	0.43%	0.78%	1.00%

Table 5: Welfare cost of inflation for different target values of buyers' surplus share

For every level of the target buyers' surplus share, the cost of inflation is *higher* with choice. As discussed in Section 8.3.3, this is due to both the greater effect of inflation on seller entry and the stronger impact of entry on the expected match surplus – due to both quantity and quality effects. Also, the cost of inflation increases when we increase the target buyers' surplus share, $\theta(n)$. Interestingly, this is the opposite to what we observe in models featuring Nash bargaining, such as Lagos and Wright (2005), where the cost of inflation is decreasing in buyers' bargaining power.

Table 6 reports the welfare losses at the Friedman rule for the same exercise.

	$\theta(n) = 0.4$	$\theta(n) = 0.5$	$\theta(n) = 0.6$
Choice			
Welfare loss at FR vs efficient	0.33%	0.16%	0.04%
No choice			
Welfare loss at FR vs efficient	0.45%	0.47%	0.33%

Table 6: Welfare loss at Friedman rule for different target values of buyers' surplus share

The welfare loss at the Friedman rule is always *lower* with choice than without choice for every level of the target buyers' surplus share. As discussed in Section 8.3.4, seller over-entry with choice leads to a higher average quality at the Friedman rule than the efficient outcome, which tends to alleviate the extent of the welfare loss.

With choice, the welfare loss at the Friedman rule *decreases* when we increase the target buyers' surplus share. Intuitively, this is because greater choice among sellers gives buyers greater effective "bargaining power", so the seller-buyer ratio n is higher when we target a higher buyers' surplus share. We know from Section 6.1.1 that as n becomes large (i.e. as $k \to 0$), we approach the efficient allocation and thus the welfare loss at the Friedman rule decreases. Without choice, the reasoning in Section 6.1.1 does not apply as changes in n do not affect the distribution of chosen goods.

9 Conclusion

This paper introduces consumer choice into a search-theoretic model of monetary exchange and examines its implications for efficiency and the welfare cost of inflation. Consumers can simultaneously meet a number of different sellers and *choose* the seller with whom they wish to trade. Consumer choice is influenced by seller-specific random utility shocks that are private information for the buyer. We allow for seller entry, which is important because greater seller entry implies greater choice for buyers. We focus on competitive search equilibrium and show that the Friedman rule does not generally deliver the efficient allocation in our environment.

In our baseline calibration, there is underconsumption of all goods and under-entry of sellers. At the Friedman rule, there is underconsumption and over-entry of sellers. We find that choice matters quantitatively. The welfare cost of increasing inflation from 0% to 10% is more than twice as high when we incorporate consumer choice: around 1.5% of consumption versus 0.6% without choice. Intuitively, higher inflation has a greater impact on the average match surplus when there is consumer choice due to its greater effect on seller entry, which affects both the average quantity traded *and* the average quality of chosen goods.

As shown in Lagos and Wright (2005), a deeper understanding of the microfoundations of monetary exchange is essential for analyzing the effects of monetary policy. Our paper extends this fundamental insight in a new direction. We show that modelling the process of consumer choice as part of this microfoundation yields important insights regarding both the cost of inflation and the nature of the inefficiencies that may arise in markets with monetary exchange. Our results suggest that incorporating choice into models of monetary exchange is both qualitatively and quantitatively important for understanding the effects of monetary policy. We hope this spurs further research into the implications of choice in monetary environments.

References

A. Berentsen, G. Menzio, and R. Wright. Inflation and unemployment in the long run. *American Economic Review*, 101(1):371–98, February 2011.

Z. Bethune, M. Choi, and R. Wright. Frictional goods markets: Theory and applications. *The Review of Economic Studies (forthcoming)*, 2020.

T. F. Cooley and G. D. Hansen. The inflation tax in a real business cycle model. *The American Economic Review*, pages 733–748, 1989.

S. M. Davoodalhosseini. Constrained efficiency with adverse selection and directed search. *Journal of Economic Theory*, 183:568–593, 2019.

M. Dong. Inflation and variety. *International Economic Review*, 51(2):401–420, 2010.

M. Dong and J. H. Jiang. Money and price posting under private information. *Journal of Economic Theory*, 150:740–777, 2014.

M. Faig and B. Jerez. A theory of commerce. *Journal of Economic Theory*, 122(1): 60–99, 2005.

M. Faig and B. Jerez. Inflation, prices, and information in competitive search. *Advances in Macroeconomics*, 6(1):1–34, 2006.

V. Guerrieri. Inefficient unemployment dynamics under asymmetric information. *Journal of Political Economy*, 116(4):667–708, 2008.

V. Guerrieri, R. Shimer, and R. Wright. Adverse selection in competitive search equilibrium. *Econometrica*, 78(6):1823–1862, 2010.

A. J. Hosios. On the Efficiency of Matching and Related Models of Search and Unemployment. *Review of Economic Studies*, 57(2):279–298, 1990.

R. Lagos and R. Wright. A unified framework for monetary theory and policy analysis. *Journal of Political Economy*, 113(3):463–484, 2005.

R. Lagos, G. Rocheteau, and R. Wright. Liquidity: A new monetarist perspective. *Journal of Economic Literature*, 55(2):371–440, 2017.

R. E. Lucas and J. P. Nicolini. On the stability of money demand. *Journal of Monetary Economics*, 73:48–65, 2015.

R. J. Lucas. Inflation and welfare. *Econometrica*, 68(2):247–274, 2000.

S. Mangin and B. Julien. Efficiency in search and matching models: a generalized Hosios condition. *Working paper*, 2020.

E. Maskin and J. Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984.

E. Moen. Competitive Search Equilibrium. *Journal of Political Economy*, 105(2): 385–411, 1997.

E. R. Moen and A. Rosen. Incentives in competitive search equilibrium. *The Review of Economic Studies*, 78(2):733–761, 2011.

G. Rocheteau and R. Wright. Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium. *Econometrica*, 73(1):175–202, 2005.

G. Rocheteau and R. Wright. Inflation and welfare with trading frictions. *Monetary Policy in Low Inflation Economies, ed. Ed Nosal and D. Altig*, 2009.

R. Shimer. PhD Thesis. *M.I.T.*, 1996.

N. Stokey, R. Lucas, and E. C. Prescott. *Recursive Methods in Economic Dynamics.* Harvard University Press, 1989.

R. Wright, P. Kircher, B. Julien, and V. Guerrieri. Directed search and competitive search equilibrium: A guided tour. *Journal of Economic Literature (forthcoming)*, 2020.

10 Online Appendix: Proofs

10.1 Proofs for Section 4

Proof of Lemma 1

Part 1. Using the fact that the distribution of the maximum of $N \geq 1$ draws is $(G(a))^N$, and weighting by the probability $P_N(n)$ that exactly N sellers meet a buyer, conditional on $N \geq 1$, we obtain

$$\tilde{G}(a;n) = \frac{\sum_{N=1}^{\infty} P_N(n)(G(a))^N}{\alpha(n)}. \tag{50}$$

Given that we assume a Poisson distribution, the distribution of chosen goods and the average quality of a chosen good can be derived as follows. Substituting $P_N(n) = \frac{n^N e^{-n}}{N!}$ and $\alpha(n) = 1 - e^{-n}$ into (50) yields

$$\tilde{G}(a;n) = \frac{e^{-n} \sum_{N=0}^{\infty} \frac{(nG(a))^N}{N!} - e^{-n}}{1 - e^{-n}} \tag{51}$$

which, using the fact that $\sum_{N=0}^{\infty} \frac{(nG(a))^N}{n!} = e^{-n(G(a))}$, simplifies to (4).

Part 2. For any given $n > 0$, we have $\tilde{G}(a;n) < G(a)$ for all $a \in [a_0, \bar{a}]$. To see this, let $w_N(n) = P_N(n)/\alpha(n)$. Using (50), we have $\tilde{G}(a;n) = \sum_{N=1}^{\infty} w_N(n)(G(a))^N$. Since $\tilde{G}(a;n)$ is a weighted average of $(G(a))^N$ for all $N > 1$, and $(G(a))^N < G(a)$ for all $N > 1$ and $a \in (a_0, \bar{a})$, and $G(a)^N = G(a)$ for $a = a_0$ or $a = \bar{a}$, we have $\tilde{G}(a;n) < G(a)$. Therefore, $\tilde{G}$ first order stochastically dominates G and $\tilde{a}(n) > E_G(a)$.

Part 3. Taking the limit as $n \to 0$, we have

$$\lim_{n \to 0} \tilde{G}(a;n) = \lim_{n \to 0} \frac{e^{-n(1-G(a))} - e^{-n}}{1 - e^{-n}} = G(a) \tag{52}$$

by an application of L'Hopital's rule. Therefore, $\tilde{a}(n) \to E_G(a)$.

Part 4. Taking the limit as $n \to \infty$, we have

$$\lim_{n\to\infty} \tilde{G}(a;n) = \lim_{n\to\infty} \frac{e^{-n(1-G(a))} - e^{-n}}{1-e^{-n}} = 0 \tag{53}$$

for any $a \in [a_0, \bar{a})$ and $\lim_{n\to\infty} \tilde{G}(\bar{a};n) = 1$. Therefore, we have $\tilde{a}(n) \to \bar{a}$. ■

Proof of Lemma 2

Applying Leibniz' integral rule, the derivative of $\tilde{f}(n)$ with respect to n is

$$\tilde{f}'(n) = \int_{a_0}^{\bar{a}} f(a) \frac{\partial \tilde{g}(a;n)}{\partial n} da. \tag{54}$$

First, we show that there exists a unique cutoff $\hat{a} \in [a_0, \bar{a}]$ s.t. $\frac{\partial \tilde{g}(a;n)}{\partial n} > 0$ for $a > \hat{a}$ and $\frac{\partial \tilde{g}(a;n)}{\partial n} < 0$ for $a < \hat{a}$. To start with, we have

$$\tilde{g}(a;n) = \frac{ng(a)e^{-n(1-G(a))}}{1-e^{-n}}. \tag{55}$$

Differentiating (55) with respect to n, we obtain

$$\frac{\partial \tilde{g}(a;n)}{\partial n} = g(a) \left[\frac{e^{-n(1-G(a))}[(1-n(1-G(a)))(1-e^{-n}) - ne^{-n}]}{(1-e^{-n})^2} \right] \tag{56}$$

and therefore $\frac{\partial \tilde{g}(a;n)}{\partial n} > 0$ if and only if

$$(1-n(1-G(a)))(1-e^{-n}) - ne^{-n} > 0, \tag{57}$$

or, equivalently,

$$G(a) > \frac{1}{1-e^{-n}} - \frac{1}{n}. \tag{58}$$

Defining $\hat{a} = G^{-1}\left(\frac{1}{1-e^{-n}} - \frac{1}{n}\right)$, we have $\frac{\partial \tilde{g}(a;n)}{\partial n} > 0$ if and only if $a > \hat{a}$.

We can use the cutoff $\hat{a}$ to rewrite $\tilde{f}'(n)$ as follows:

$$\tilde{f}'(n) \equiv \int_{a_0}^{\hat{a}} f(a) \frac{\partial \tilde{g}(a;n)}{\partial n} da + \int_{\hat{a}}^{\bar{a}} f(a) \frac{\partial \tilde{g}(a;n)}{\partial n} da. \tag{59}$$

We therefore have $\tilde{f}'(n) > 0$ if and only if

$$\int_{\hat{a}}^{\bar{a}} f(a) \frac{\partial \tilde{g}(a;n)}{\partial n} da > -\int_{a_0}^{\hat{a}} f(a) \frac{\partial \tilde{g}(a;n)}{\partial n} da > 0. \tag{60}$$

Given that $f'(a) > 0$, and both sides of (60) are positive, by definition of $\hat{a}$, a sufficient condition for $\tilde{f}'(n) > 0$ is

$$\int_{\hat{a}}^{\bar{a}} f(\hat{a}) \frac{\partial \tilde{g}(a;n)}{\partial n} da \geq -\int_{a_0}^{\hat{a}} f(\hat{a}) \frac{\partial \tilde{g}(a;n)}{\partial n} da, \tag{61}$$

which is true iff $\int_{\hat{a}}^{\bar{a}} \frac{\partial \tilde{g}(a;n)}{\partial n} da \geq -\int_{a_0}^{\hat{a}} \frac{\partial \tilde{g}(a;n)}{\partial n} da$, or equivalently $\int_{a_0}^{\bar{a}} \frac{\partial \tilde{g}(a;n)}{\partial n} da \geq 0$. Applying Leibniz' integral rule again, $\int_{a_0}^{\bar{a}} \frac{\partial \tilde{g}(a;n)}{\partial n} da = \frac{\partial}{\partial n} \int_{a_0}^{\bar{a}} \tilde{g}(a;n) da = 0$, since $\int_{a_0}^{\bar{a}} \tilde{g}(a;n) da = 1$. Therefore, $\tilde{f}'(n) > 0$. ■

10.2 Proofs for Section 5

Proof of Proposition 1

The first-order condition with respect to q_a is

$$\alpha(n)[au'(q_a) - c'(q_a)]\tilde{g}(a;n) = 0 \tag{62}$$

and the first order-condition with respect to n is

$$\alpha'(n)\tilde{s}(n; \{q_a\}_{a \in [a_0, \bar{a}]}) + \alpha(n)\tilde{s}'(n; \{q_a\}_{a \in [a_0, \bar{a}]}) = k. \tag{63}$$

We can verify that $s_a^* = au(q_a^*) - c(q_a^*)$ is strictly increasing in a. Differentiating s_a^*,

$$\frac{ds_a^*}{da} = u(q_a^*) + [au'(q_a^*) - c'(q_a^*)] \frac{dq^*}{da}. \tag{64}$$

Since $au'(q_a^*) - c'(q_a^*) = 0$ by (62) if $n^* > 0$, we have $\frac{ds_a^*}{da} = u(q_a^*) > 0$ for all $a \in (a_0, \bar{a}]$. Given that s_a^* is strictly increasing in a and $s_0^* \geq 0$ where $s_0^* \equiv a_0 u(q_0) - c(q_0)$, we have $s_a^* \geq 0$ for all $a \in [a_0, \bar{a}]$. Therefore, all chosen goods $a \in [a_0, \bar{a}]$ are traded if $a_0 > 0$, and q_a satisfies $au'(q_a) = c'(q_a)$. If $a_0 = 0$, we have $q_a = 0$ since $\lim_{q \to 0} c'(q)/u'(q) = 0$.

Since s_a^* is strictly increasing in a, the planner wants the good with the highest utility shock a to be chosen in every meeting. The distribution of chosen goods,

$\tilde{G}(a; n)$, is therefore equal to distribution of the maximum utility shock drawn by the buyer, given by (4).

Existence and uniqueness of the solution to the planner's problem follows from Corollary 1, which follows from Proposition 2, both of which are proven below. For the planner's problem, we know that $s_a^* \geq 0$ for all $a \in [a_0, \bar{a}]$ and thus all chosen goods are traded. Setting $i = 0$ in Corollary 1 results in equilibrium conditions that are equivalent to the planner's FOCs. It follows that there exists a unique solution to the planner's problem with $n^* > 0$ provided that Assumption 3 holds, except that $q_a^0 = q_a^*$ since q_a^* does not depend directly on n. That is, Assumption 2 suffices. ■

10.3 Proof of Proposition 2

Our strategy is to solve for the equilibrium in two stages. First, we take z and n as given and solve for $\{(q_a, d_a)\}_{a \in [a_0, \bar{a}]}$ (inner maximization problem). Second, we solve for z and n (outer maximization problem) given the solutions for $\{(q_a, d_a)\}_{a \in [a_0, \bar{a}]}$.

We first solve the inner and outer maximization problems. Next, we use the results to prove Parts 1 to 8 of Proposition 2. Finally, we prove existence and uniqueness of equilibrium. Proofs for all lemmas used in this section to prove Proposition 2 are found at the end of this section (unless included earlier).

Stage 1. Inner maximization problem

In the first stage, taking $z > 0$ and $n > 0$ as given (we later prove this), the market makers' problem is to maximize (21) subject to (22) at equality, plus a liquidity constraint $d_a \leq z$ for all $a \in [a_0, \bar{a}]$, the IC constraint (24), and the IR constraint (23). Ignoring constants, the market maker's inner maximization problem is:

$$\max_{\{(q_a, d_a)\}_{a \in [a_0, \bar{a}]}} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \left[a u(q_a) - \frac{d_a}{\gamma} \right] d\tilde{G}(a; n) - i \frac{z}{\gamma} \right\}, \tag{65}$$

subject to

$$\frac{\alpha(n)}{n} \int_{a_0}^{\bar{a}} \left[-c(q_a) + \frac{d_a}{\gamma} \right] d\tilde{G}(a; n) = k, \tag{66}$$

and, for all $a, a' \in [a_0, \bar{a}]$,

$$
\begin{aligned}
(67) \qquad d_a &\leq z, \\
(68) \qquad au(q_a) - \frac{d_a}{\gamma} &\geq au(q_{a'}) - \frac{d_{a'}}{\gamma}, \\
(69) \qquad au(q_a) - \frac{d_a}{\gamma} &\geq 0, \\
(70) \qquad d_a, q_a &\geq 0.
\end{aligned}
$$

To solve the inner maximization problem (65), we transform the above problem as follows. Defining $v_a \equiv au(q_a) - d_a/\gamma$, the buyer's ex post trading surplus, and $\dot{v}_a \equiv v'(a)$, the following lemma simplifies the (IC) constraint. This is a standard result and the proof is omitted.

Lemma 5. *The incentive compatibility (IC) constraint holds if and only if (i) $q'(a) \geq 0$, and (ii) $\dot{v}_a = u(q_a)$.*

We can now use $v_a \equiv au(q_a) - d_a/\gamma$ and Lemma 5 to re-write the problem as an optimal control problem where q_a is the control variable, v_a is the state variable, and δ is the Lagrange multiplier associated with the seller entry constraint (66). For simplicity, we assume that $a_0 = 0$.

In the first stage, we take z, n, δ as given and later solve for these. Given that $a_0 = 0$, we have $v_0 = 0$. Using $v_a \equiv au(q_a) - d_a/\gamma$ to eliminate d_a in the above, and substituting in the constraint (66), the inner maximization problem becomes

$$
(71) \qquad \max_{\{(q_a, v_a)\}_{a \in [a_0, \bar{a}]}} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \{(1-\delta)v_a + \delta \left[au(q_a) - c(q_a)\right]\} \tilde{g}(a; n) da - \delta nk - i\frac{z}{\gamma} \right\},
$$

subject to $v_0 = 0$ and, for all $a \in [a_0, \bar{a}]$,

$$
\begin{aligned}
(72) \qquad au(q_a) - v_a &\leq \frac{z}{\gamma}, \\
(73) \qquad \dot{v}_a &= u(q_a), \\
(74) \qquad q'(a) &\geq 0, \\
(75) \qquad q_a, v_a &\geq 0.
\end{aligned}
$$

The inner maximization problem is a standard optimal control problem with q_a as the control variable and v_a as the state variable. We can therefore apply the Maximum

Principle to find the necessary conditions for the optimal path of the control and state variables. To solve the inner maximization problem, we ignore the condition $q'(a) \geq 0$ and later verify that it holds in Lemma 9. Ignoring the constants, the current value Hamiltonian for the optimal control problem is:

$$H = \alpha(n)\{(1-\delta)v_a + \delta\left[au(q_a) - c(q_a)\right]\}\tilde{g}(a;n) + \lambda_a u(q_a) \tag{76}$$

where λ_a is the costate variable, and the Lagrangian is:

$$L = \alpha(n)\{(1-\delta)v_a + \delta\left[au(q_a) - c(q_a)\right]\}\tilde{g}(a;n) + \lambda_a u(q_a) + \mu_a\left[\frac{z}{\gamma} - au(q_a) + v_a\right] + \theta_a q_a + \eta_a v_a \tag{77}$$

where μ_a, θ_a and η_a are the Lagrangian multipliers associated with the liquidity constraint, non-negativity constraint, and IR constraint respectively.

The FOCs and the transversality condition are as follows:

$$\frac{\partial L}{\partial q_a} = \alpha(n)\delta\left[au'(q_a) - c'(q_a)\right]\tilde{g}(a;n) + (\lambda_a - \mu_a a)\,u'(q_a) + \theta_a = 0, \tag{78}$$

$$\frac{\partial L}{\partial v_a} = (1-\delta)\alpha(n)\tilde{g}(a;n) + \mu_a + \eta_a = -\dot{\lambda}_a, \tag{79}$$

$$\frac{\partial L}{\partial \lambda_a} = \dot{v}_a = u(q_a), \tag{80}$$

$$\lambda_{\bar{a}} v_{\bar{a}} = 0. \tag{81}$$

For the inequality constraints, the conditions are:

$$\mu_a \geq 0,\ \mu_a\left(\frac{z}{\gamma} - au(q_a) + v_a\right) = 0, \tag{82}$$

$$\theta_a \geq 0,\ \theta_a q_a = 0, \tag{83}$$

$$\eta_a \geq 0,\ \eta_a v_a = 0. \tag{84}$$

The following lemma provides expressions for λ_a and Σ_{a_c}, where $\Sigma_a \equiv \int_a^{\bar{a}} \mu_x dx$.

Lemma 6. *For all $a \in [a_0, a_c]$, we have the following:*

$$\lambda_a = \alpha(n)(1-\delta)[1-\tilde{G}(a;n)] + \Sigma_{a_c} + \int_a^{\bar{a}} \eta_x dx \tag{85}$$

and

$$\Sigma_{a_c} = \frac{\alpha(n)}{\bar{a}} \int_{a_c}^{\bar{a}} [\delta(x-a_c)\tilde{g}(x;n) + (1-\delta)(\tilde{G}(a_c;n) - \tilde{G}(x;n)]dx. \tag{86}$$

By Lemma 3, there are three intervals to consider.

Case 1. For any $a \in [a_0, a_b]$, $v_a = 0$ for all a and therefore $q_a = 0$.

Case 2. For any $a \in [a_b, a_c]$, we have $\theta_a = 0$ and $\mu_a = 0$, so q_a solves

$$\alpha(n)\delta\left[au'(q_a) - c'(q_a)\right]\tilde{g}(a;n) \ = -\lambda_a u'(q_a). \tag{87}$$

Using expression (85) for λ_a from Lemma 6, we can write the above as

$$(a - \phi(a))u'(q_a) = c'(q_a) \tag{88}$$

where

$$\phi(a) = -\left(\frac{1-\delta}{\delta}\right)\left(\frac{1-\tilde{G}(a;n)}{\tilde{g}(a;n)}\right) - \frac{\Sigma_{a_c}}{\alpha(n)\delta\tilde{g}(a;n)}. \tag{89}$$

Case 3. For any $a \in [a_c, \bar{a}]$, we have $\theta_a = 0$ and $q_a = q_{a_c}$ by Lemma 3.

The following two lemmas will prove useful in deriving Proposition 2.

Lemma 7. *We have $a = \phi(a)$ for all $a \leq a_b$ where $\phi(a)$ is given by (89).*

Proof. For $a = a_b$, we have $q_{a_b} = 0$. Using (88) above, we have

$$\lim_{a \to a_b} (a_b - \phi(a_b)) = \lim_{q \to 0} \frac{c'(q)}{u'(q)} = 0 \tag{90}$$

since we have $\lim_{q\to 0} \frac{c'(q)}{u'(q)} = 0$ by assumption. Therefore, by continuity of the function q_a, we have $a_b = \phi(a_b)$. Similarly, $a = \phi(a)$ for all $a < a_b$. ■

The next lemma uses our assumption that $a_0 = 0$.

Lemma 8. *If* $a_0 = 0$*, we obtain the following:*

$$\delta = 1 + \frac{\Sigma_{a_c} + \int_{a_0}^{\bar{a}} \eta_x dx}{\alpha(n)}. \tag{91}$$

To determine q_a for all $a \in [a_0, \bar{a}]$, it remains only to determine δ, a_b, and a_c.

Stage 2. Outer maximization problem

The outer maximization problem we solve next is

$$\max_{z,n,\delta} \left\{ J(n,z,\delta) - \delta n k - i\frac{z}{\gamma} \right\}, \tag{92}$$

where we define

$$J(n,z,\delta) \equiv \max_{\{(q_a,v_a)\}_{a\in[a_0,\bar{a}]}} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \{(1-\delta)v_a + \delta\left[au(q_a) - c(q_a)\right]\} \tilde{g}(a;n)da \right\}, \tag{93}$$

subject to $v_0 = 0$ and, for all $a \in [a_0, \bar{a}]$, constraints (72), (73), (74), and (75).

To solve the outer maximization problem, the function $J(n,z,\delta)$ is equivalent to

$$J(n,z,\delta) = \max_{\{(q_a,v_a)\}_{a\in[a_0,\bar{a}]}} \left\{ \begin{array}{c} \int_{a_0}^{\bar{a}} \alpha(n)\{(1-\delta)v_a + \delta\left[au(q_a) - c(q_a)\right]\}\tilde{g}(a;n)da \\ + \int_{a_0}^{\bar{a}} \left[\mu_a\left(\frac{z}{\gamma} - au(q_a) + v_a\right) + \eta_a v_a + \lambda_a u(q_a) + \theta_a q_a\right] da \end{array} \right\}. \tag{94}$$

Define $\tilde{s}(n) \equiv \int_{a_0}^{\bar{a}} s_a d\tilde{G}(a;n)$ and $\tilde{v}(n) \equiv \int_{a_0}^{\bar{a}} v_a d\tilde{G}(a;n)$. Returning to our original formulation to eliminate δ, the problem is equivalent to

$$\max_{z,n} \left\{ \hat{J}(n,z) - i\frac{z}{\gamma} \right\}, \tag{95}$$

where

$$\hat{J}(n,z) = \max_{\{(q_a,v_a)\}_{a\in[a_0,\bar{a}]}} \left\{ \alpha(n)\tilde{v}(n) + \int_{a_0}^{\bar{a}} \left[\mu_a\left(\frac{z}{\gamma} - au(q_a) + v_a\right) + \eta_a v_a + \lambda_a u(q_a) + \theta_a q_a\right] da \right\} \tag{96}$$

subject to the constraint

$$\frac{\alpha(n)}{n}[\tilde{s}(n) - \tilde{v}(n)] \leq k \tag{97}$$

and $n \geq 0$ with complementary slackness.

Using the envelope theorem, the first-order conditions for z and n respectively are

(98) $$\int_{a_0}^{\bar{a}} \mu_a da = i$$

and

(99) $$\alpha'(n)\tilde{v}(n) + \alpha(n)\tilde{v}'(n) = 0.$$

Using the fact that $\mu_a = 0$ for all $a < a_c$, by definition of a_c, we have $\int_{a_0}^{\bar{a}} \mu_a da = \Sigma_{a_c}$. The FOC for z given by (98) thus becomes:

(100) $$\Sigma_{a_c} = i,$$

Substituting $\Sigma_{a_c} = i$ into expression (91) in Lemma 8, the above yields

(101) $$\delta = 1 + \frac{i + \int_{a_0}^{\bar{a}} \eta_x dx}{\alpha(n)}.$$

Finally, we verify that the condition $q'(a) \geq 0$ is indeed satisfied. Since we assume in Assumption 1 that $G''(a) \leq 0$, Lemma 9 implies that $q'(a) \geq 0$.

Lemma 9. *If $G''(a) \leq 0$ for all $a \in [a_0, \bar{a}]$, then $q(.)$ is weakly increasing for all $a \in [a_0, \bar{a}]$ and $q'(a) > 0$ for all $a \in (a_b, a_c)$.*

Proof of Parts 1 to 8

Part 1. Follows from the definition of a_b.

Part 2. From above, for any $a \in [a_b, a_c]$, we have

(102) $$(a - \phi(a))u'(q_a) = c'(q_a)$$

where, using $\Sigma_{a_c} = i$, we have

(103) $$\phi(a) = -\left(\frac{1-\delta}{\delta}\right)\left(\frac{1-\tilde{G}(a;n)}{\tilde{g}(a;n)}\right) - \frac{i}{\alpha(n)\delta\tilde{g}(a;n)}.$$

The expression for δ can be derived as follows. Using the fact that $a_b = \phi(a_b)$

from Lemma 7, plus expression (89) for $\phi(a)$, as well as $\Sigma_{a_c} = i$, we obtain:

$$i = -\alpha(n)[\delta a_b \tilde{g}(a_b; n) + (1-\delta)(1 - \tilde{G}(a_b; n))]. \tag{104}$$

Rearranging, δ is given by the following expression:

$$\delta = \frac{1 - \tilde{G}(a_b; n) + \frac{i}{\alpha(n)}}{1 - \tilde{G}(a_b; n) - a_b \tilde{g}(a_b; n)} \tag{105}$$

which is equivalent to (30) using expression (26).

Also, $\dot{v}_a = u(q_a)$ implies $v_a - v_0 = \int_{a_0}^{a} u(q_x)dx$, so $v_a = \int_{a_0}^{a} u(q_x)dx$ since $v_0 = 0$. We can derive d_a/γ from v_a using the fact that $v_a \equiv au(q_a) - d_a/\gamma$.

Part 3. Clear from Lemma 3.

Part 4. Using $\Sigma_{a_c} = i$ and expression (86) for Σ_{a_c}, the value of a_c is given by (31).

Part 5. Clear from the definition of a_c.

Part 6. The first-order condition for $n > 0$ given by (99) can be written as

$$\alpha'(n)\tilde{s}(n) + \alpha(n)\tilde{s}'(n) = k, \tag{106}$$

using the ZPC constraint (97) at equality. More precisely, this is equivalent to

$$\alpha'(n)\tilde{s}(n; \{q_a\}_{a \in [a_0, \bar{a}]}) + \alpha(n)\tilde{s}'(n; \{q_a\}_{a \in [a_0, \bar{a}]}) = k. \tag{107}$$

The fact that n is strictly decreasing in k is proven in Lemma 12 below.

Part 7. The zero profit condition is given by (97), using the definition of v_a.

Part 8. Since v_a is increasing in a, the highest draw is always chosen and therefore the distribution of chosen goods is the distribution of the maximum given by (4). ■

Proof of existence and uniqueness

We first prove existence and uniqueness of the solution to the inner maximization problem and then prove the same for the outer maximization problem.

Inner maximization. We prove that, given z and n from the outer maximization problem, the solution to the inner maximization problem exists and is unique.

Existence. A solution to the problem exists because the set of admissible paths is non-empty and compact, and there exists an admissible path for which the objective is finite. For example, the path $q_a = 0$ and $v_a = (a-1)u(q_a)$ for all $a \in [a_0, \bar{a}]$ is admissible (since $v_0 = 0$, $au(q_a) - v_a \leq z/\gamma$, $q_a \geq 0$, $v_a \geq 0$, and $\dot{v}_a = u(q_a) + (a-1)u'(q_a)q'(a) = u(q_a)$, and $q'(a) \geq 0$). Also, the objective is finite under this path. Finally, the set of feasible paths is compact since $q_a \in [0, q^*_{\bar{a}}]$ where $q^*_{\bar{a}}$ solves $\bar{a}u'(q_{\bar{a}}) = c'(q_{\bar{a}})$ and $v_a \in [0, v_{\bar{a}}]$ where $v_{\bar{a}} = u(q^*_{\bar{a}})[\bar{a} - a_0]$ since $v_a = \int_{a_0}^{a} u(q_x)dx$.

Uniqueness. The Hamiltonian $H(q_a, v_a, \lambda_a)$, where λ_a is the co-state variable given by the Maximum Principle, is strictly concave in the control and state variables (q_a, v_a) for all a. Therefore, the solution is an optimum that solves the inner maximization problem and it is unique. To establish strict concavity, differentiating $H(q_a, v_a, \lambda_a)$ with respect to q_a yields

$$\begin{aligned} \frac{\partial H}{\partial q_a} &= \alpha(n)\delta[u'(q_a) - c'(q_a)]\tilde{g}(a;n) + \lambda_a u'(q_a), \\ \frac{\partial^2 H}{\partial q_a^2} &= \alpha(n)\delta[u''(q_a) - c''(q_a)]\tilde{g}(a;n) + \lambda_a u''(q_a) \equiv -X, \end{aligned}$$

where $X > 0$, since $u''(q_a) < 0$ and $c''(q_a) > 0$. Differentiating $H(q_a, v_a, \lambda_a)$ with respect to v_a, we obtain $\frac{\partial H}{\partial v_a} = \alpha(n)(1-\delta)\tilde{g}(a;n)$ and $\frac{\partial^2 H}{\partial v_a^2} = 0$. Finally, $\frac{\partial^2 H}{\partial v_a \partial q_a} = 0$, so we get the Hessian matrix, $\mathbb{H} = \begin{bmatrix} -X & 0 \\ 0 & 0 \end{bmatrix}$. Since $\mathbf{x}^T \mathbb{H} \mathbf{x} < 0$ for all $\mathbf{x} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$, the Hessian $\mathbb{H}$ is negative definite and the Hamiltonian is strictly concave in (q_a, v_a).

Outer maximization. We prove that, given $\{(q_a, v_a)\}_{a \in [a_0, \bar{a}]}$ from the inner maximization problem, the solution (n, z) to the outer maximization problem exists and is unique, and n, z are interior solutions with $n, z > 0$ provided that Assumption 3 holds. To establish this result, we first prove that there exists a non-empty set of solutions n, denoted by $N(k)$, that solves the problem. We then show that equilibrium is unique if $n > 0$ for all $n \in N(k)$, and finally we prove that $n > 0$ for any $n \in N(k)$.

Taking $\{(q_a, v_a)\}_{a \in [a_0, \bar{a}]}$ as given by the inner maximization problem, and ignoring constants, the outer maximization problem is equivalent to

$$\max_{z,n} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \left[au(q_a) - \frac{d_a}{\gamma} \right] d\tilde{G}(a;n) + (\Sigma_{a_c} - i) \frac{z}{\gamma} \right\}, \tag{108}$$

subject to

$$\frac{\alpha(n)}{n}\int_{a_0}^{\bar{a}}\left[-c(q_a)+\frac{d_a}{\gamma}\right]d\tilde{G}(a;n)\leq k \tag{109}$$

and $n \geq 0$ with complementary slackness, where $\{(q_a, v_a)\}_{a\in[a_0,\bar{a}]}$ solves the inner maximization problem.

Lemma 10. *The set of solutions $N(k)$ is nonempty and upper hemicontinuous.*

Proof. Since $\alpha(n)$ is a bijection, we can rewrite (108) in terms of α as follows:

$$\max_{z,\alpha}\left\{\alpha\int_{a_0}^{\bar{a}}\left[au(q_a)-\frac{d_a}{\gamma}\right]d\tilde{G}(a;\alpha)+(\Sigma_{a_c}-i)\,\frac{z}{\gamma}\right\}. \tag{110}$$

The objective function is continuous and, without loss of generality, we can restrict (z, α) to the following compact set:

$$\Delta = \{(z,\alpha) : \alpha \in [0,1],\ z/\gamma \in [0, \bar{a}u(q_{\bar{a}})]\} \tag{111}$$

since $q \in [0, q^*_{\bar{a}}]$ where $q^*_{\bar{a}}$ solves $\bar{a}u'(q_{\bar{a}}) = c'(q_{\bar{a}})$, and we have $z/\gamma < \bar{a}u(q_{\bar{a}})$. The constraint (109) can therefore be written as $(z, \alpha) \in \Gamma(k)$ for all $k \geq 0$, where $\Gamma(k)$ is a continuous and compact-valued correspondence. Applying the Theorem of the Maximum (Theorem 3.6 in Stokey, Lucas, and Prescott, 1989), the correspondence that gives the set of solutions for α is nonempty and upper hemicontinuous, and therefore also $N(k)$ is nonempty and upper hemicontinuous. ■

The following lemma establishes that any strictly positive solution $n \in N(k)$ must be unique. Since we know that $z = d_{a_c} > 0$ where $d_a/\gamma = au(q_a) - v_a$, and $\{(q_a, v_a)\}_{a\in[a_0,\bar{a}]}$ is given by the inner maximization problem, Lemma 11 implies that any solution (n, z) where $n > 0$ is unique.

Lemma 11. *If $N^+ \subseteq N(k)$ and $N^+ \subseteq \mathbb{R}_+\backslash\{0\}$, then $N^+ = \{n\}$.*

Proof. Consider any solution $n \in N(k)$ such that $n > 0$. Defining $\Phi(n) = \alpha(n)\tilde{v}(n)$, the solutions n satisfy the first-order condition (99), which says $\Phi'(n) = 0$. We show that $\Phi''(n) < 0$ and thus any solution is unique. Using (55), we have

$$\Phi(n) = \int_{a_0}^{\bar{a}} ne^{-n(1-G(a))}v(a)g(a)da. \tag{112}$$

Using Leibniz's integral rule, plus the envelope theorem,

$$\Phi'(n) = \int_{a_0}^{\bar{a}} e^{-n(1-G(a))} v(a) g(a) da - \int_{a_0}^{\bar{a}} n(1-G(a)) e^{-n(1-G(a))} v(a) g(a) da. \tag{113}$$

By integration by parts on the right integral of (113), we obtain

$$\Phi'(n) = \int_{a_0}^{\bar{a}} e^{-n(1-G(a))} (1-G(a)) v'(a) da + e^{-n} v(a_0) > 0. \tag{114}$$

Differentiating (114), we find that

$$\Phi''(n) = -\left(\int_{a_0}^{\bar{a}} e^{-n(1-G(a))} (1-G(a))^2 v'(a) da + e^{-n} v(a_0) \right) < 0. \tag{115}$$

The fact that $\Phi''(n) < 0$ follows from the fact that $v'(a) = u(q_a) \geq 0$ for all a and $v'(a) > 0$ for some a and also $v(a_0) = 0$. Therefore, any solution $n > 0$ is unique. ■

From Lemma 10, we know that, for any given $k \geq 0$, there exists a non-empty set of solutions $N(k)$ that solves problem (108). We also know that any solution z is interior, since $z/\gamma = \bar{a}u(q_{\bar{a}})$ implies $v_{\bar{a}} = \bar{a}u(q_{\bar{a}}) - \bar{z}/\gamma = 0$ and therefore $v_a = 0$ for all $a \in [a_0, \bar{a}]$. We now prove that, for any $n \in N(k)$, we have $n \in \mathbb{R}_+ \backslash \{0\}$ provided that Assumption 3 holds. Also, the function $n(k)$ is strictly decreasing in k.

Lemma 12. *Any solution $n \in N(k)$ is interior, i.e. $n \in \mathbb{R}_+ \backslash \{0\}$. The function $n(k)$ is strictly decreasing in k.*

Proof. First, we show there exists an interior solution $n > 0$. Define $\Lambda(n) \equiv \alpha(n)\tilde{s}(n)$. The first-order condition (106) says $\Lambda'(n) = k$. We prove there exists $n > 0$ such that $\Lambda'(n) = k$ if Assumption 3 holds. We have $\lim_{n\to\infty} \Lambda'(n) = 0$, and

$$\lim_{n\to 0} \Lambda'(n) = \int_{a_0}^{\bar{a}} \lim_{n\to 0} s(a; q_a(n)) dG(a) \tag{116}$$

where $\lim_{n\to 0} s(a; q_a(n)) = s(a; \lim_{n\to 0} q_a(n))$. If the following condition holds:

$$E_G[au(q_a^0) - c(q_a^0)] > k \tag{117}$$

where $q_a^0 \equiv \lim_{n\to 0} q_a(n)$, there exists $n > 0$ that satisfies $\Lambda'(n) = k$ provided that $\Lambda''(n) < 0$ (which we prove below).

Next, any interior solution $n > 0$ is better than $n = 0$. Define the value function:

$$V(k,\gamma) \equiv \max_{z,n} \left\{ \alpha(n) \int_{a_0}^{\bar{a}} \left[au(q_a) - \frac{d_a}{\gamma} \right] d\tilde{G}(a;n) + (\Sigma_{a_c} - i)\frac{z}{\gamma} \right\}. \tag{118}$$

Since we know that z is interior, we have $V(k,\gamma) \equiv \max_n \{\alpha(n)\tilde{v}(n)\}$ since $\int_{a_0}^{\bar{a}} \mu_a = i$. If $n = 0$ then $V(k,\gamma) = 0$. If $n > 0$, $V(k,\gamma) \equiv \max_n \{\alpha(n)\tilde{s}(n) - nk\}$ using constraint (109) with equality. Letting $\Lambda(n) = \alpha(n)\tilde{s}(n)$, we have $V(k,\gamma) > 0$ provided that $\Lambda(n) - nk > 0$. Thus the candidate solution $n > 0$ is better than $n = 0$ provided that $\Lambda(n) > nk$ for $n > 0$. Using the fact that $\Lambda'(n) = k$, it suffices to show that $\Lambda''(n) < 0$ and $\frac{\Lambda'(n)n}{\Lambda(n)} < 1$ for $n > 0$. Similarly to Lemma 11, starting with

$$\Lambda(n) = \int_{a_0}^{\bar{a}} ne^{-n(1-G(a))} s(a)g(a)da, \tag{119}$$

and using Leibniz's integral rule, plus the envelope theorem, yields

$$\Lambda'(n) = \int_{a_0}^{\bar{a}} e^{-n(1-G(a))} s(a)g(a)da - \int_{a_0}^{\bar{a}} n(1-G(a))e^{-n(1-G(a))} s(a)g(a)da. \tag{120}$$

Therefore, we have

$$\frac{\Lambda'(n)n}{\Lambda(n)} = \frac{\int_{a_0}^{\bar{a}} ne^{-n(1-G(a))} s(a)g(a)da - \int_{a_0}^{\bar{a}} n^2(1-G(a))e^{-n(1-G(a))} s(a)g(a)da}{\int_{a_0}^{\bar{a}} ne^{-n(1-G(a))} s(a)g(a)da} < 1. \tag{121}$$

Finally, we have $\Phi(n) = \Lambda(n) - nk$ for $n > 0$, so $\Phi'(n) = \Lambda'(n) - k$ and $\Phi''(n) = \Lambda''(n)$. Since $\Phi''(n) < 0$ from the proof of Lemma 11, we have $\Lambda''(n) < 0$. It follows that, for any $n \in N(k)$, we have $n > 0$. Since we assume $k > 0$, this implies $n \in \mathbb{R}_+ \backslash \{0\}$.

Since n is unique by Lemma 11, there is a function $n : \mathbb{R}_+ \backslash \{0\} \to \mathbb{R}_+ \backslash \{0\}$ such that $n(k)$ solves $\Lambda'(n) = k$. Clearly, n is strictly decreasing in k since $\Lambda''(n) < 0$. ■

Proof of Lemma 6

Start with the fact that

$$(1-\delta)\alpha(n)\tilde{g}(a;n) + \mu_a + \eta_a = -\dot{\lambda}_a \tag{122}$$

from the FOC (79) above. Integrating both sides over $[a, \bar{a}]$, we obtain

$$(123) \qquad -\int_a^{\bar{a}} \dot{\lambda}_x dx = \int_a^{\bar{a}} (1-\delta)\alpha(n)\tilde{g}(x;n)dx + \int_a^{\bar{a}} \mu_x dx + \int_a^{\bar{a}} \eta_x dx$$

and therefore

$$(124) \qquad -(\lambda_{\bar{a}} - \lambda_a) = \alpha(n)(1-\delta)\int_a^{\bar{a}} \tilde{g}(x;n)dx + \int_a^{\bar{a}} \mu_x dx + \int_a^{\bar{a}} \eta_x dx.$$

The transversality condition $\lambda_{\bar{a}} v_{\bar{a}} = 0$ implies $\lambda_{\bar{a}} = 0$ since $v_{\bar{a}} > 0$. Substituting $\Sigma_a \equiv \int_a^{\bar{a}} \mu_x dx$ into the above, and setting $\lambda_{\bar{a}} = 0$ yields

$$(125) \qquad \lambda_a = \alpha(n)(1-\delta)\int_a^{\bar{a}} \tilde{g}(x;n)dx + \Sigma_a + \int_a^{\bar{a}} \eta_x dx.$$

Now, $\mu_a = 0$ for all $a \in [a_0, a_c]$, thus $\Sigma_a = \int_a^{\bar{a}} \mu_x dx = \int_{a_c}^{\bar{a}} \mu_x dx = \Sigma_{a_c}$ for all $a \in [a_0, a_c]$. Substituting into (125), and using the fact that $\int_a^{\bar{a}} \tilde{g}(x;n)dx = [\tilde{G}(x;n)]_a^{\bar{a}} = 1 - \tilde{G}(a;n)$, we obtain (85).

For the second part, using (78) and Lemma 3, for all $a \in [a_c, \bar{a}]$ we have

$$(126) \qquad \alpha(n)\delta\left[au'(\bar{q}) - c'(\bar{q})\right]\tilde{g}(a;n) + \left(\lambda_a - \mu_a a\right)u'(\bar{q}) = 0$$

where $\bar{q} \equiv q_{a_c}$, and, for all $a \in [a_c, \bar{a}]$, we also have

$$(127) \qquad \alpha(n)\delta\left[a_c u'(\bar{q}) - c'(\bar{q})\right]\tilde{g}(a;n) + \lambda_{a_c} u'(\bar{q}) = 0.$$

Using the above two equations, and dividing both sides by $u'(\bar{q})$, we obtain

$$(128) \qquad \alpha(n)\delta(a - a_c)\tilde{g}(a;n) = -\lambda_a + \mu_a a + \lambda_{a_c}.$$

Substituting (125) for both λ_a and λ_{a_c} into the above, and simplifying, yields

$$(129) \qquad \alpha(n)[\delta(a-a_c)\tilde{g}(a;n) + (1-\delta)(\tilde{G}(a_c;n) - \tilde{G}(a;n)] = -\Sigma_a + \mu_a a + \Sigma_{a_c}.$$

Finally, $\Sigma_a = \int_a^{\bar{a}} \mu_x dx$ implies that $\dot{\Sigma}_a = -\mu_a$ and thus we obtain

$$(130) \qquad \alpha(n)[\delta(a-a_c)\tilde{g}(a;n) + (1-\delta)(\tilde{G}(a_c;n) - \tilde{G}(a;n)] = -\Sigma_a - \dot{\Sigma}_a a + \Sigma_{a_c}.$$

Integrating both sides over $[a_c, \bar{a}]$, we have

(131)
$$\alpha(n)\int_{a_c}^{\bar{a}}[\delta(x-a_c)\tilde{g}(x;n)+(1-\delta)(\tilde{G}(a_c;n)-\tilde{G}(x;n)]dx = \int_{a_c}^{\bar{a}}\left(-\Sigma_x - \dot{\Sigma}_x x + \Sigma_{a_c}\right)dx$$

where $\int_{a_c}^{\bar{a}}\left(-\Sigma_x - \dot{\Sigma}_x x + \Sigma_{a_c}\right)dx = -\left(\int_{a_c}^{\bar{a}}\Sigma_x + \dot{\Sigma}_x x\ dx\right) + [\Sigma_{a_c}x]_{a_c}^{\bar{a}}$. Using integration by parts, $\int_{a_c}^{\bar{a}}\Sigma_x + \dot{\Sigma}_x x\ dx = [\Sigma_x x]_{a_c}^{\bar{a}} = \Sigma_{\bar{a}}\bar{a} - \Sigma_{a_c}a_c = -\Sigma_{a_c}a_c$, and $[\Sigma_{a_c}x]_{a_c}^{\bar{a}} = \Sigma_{a_c}\bar{a} - \Sigma_{a_c}a_c$. Substituting $\int_{a_c}^{\bar{a}}\left(-\Sigma_x - \dot{\Sigma}_x x + \Sigma_{a_c}\right)dx = \Sigma_{a_c}\bar{a}$ into the above gives us

$$\alpha(n)\int_{a_c}^{\bar{a}}[\delta(x-a_c)\tilde{g}(x;n)+(1-\delta)(\tilde{G}(a_c;n)-\tilde{G}(x;n)]dx = \Sigma_{a_c}\bar{a} \tag{132}$$

and we therefore obtain (86). ■

Proof of Lemma 8

To start with, we have

$$\alpha(n)\delta\left[au'(q_a)-c'(q_a)\right]\tilde{g}(a;n) + (\lambda_a - \mu_a a)\,u'(q_a) + \theta_a = 0 \tag{133}$$

from the FOC (78) for q_a. Dividing both sides by q_a, we obtain

$$\alpha(n)\delta\left[a - \frac{c'(q_a)}{u'(q_a)}\right]\tilde{g}(a;n) + (\lambda_a - \mu_a a) = \frac{-\theta_a}{u'(q_a)}. \tag{134}$$

Taking the limit as $q_a \to 0$, and using $\lim_{q\to 0} u'(q) = +\infty$ and $\lim_{q\to 0}\frac{c'(q)}{u'(q)} = 0$ yields

(135)
$$\lim_{q\to 0}\alpha(n)\delta\left[a - \frac{c'(q)}{u'(q)}\right]\tilde{g}(a;n) + (\lambda_a - \mu_a a) + \frac{\theta_a}{u'(q)} = \alpha(n)\delta a\tilde{g}(a;n) + (\lambda_a - \mu_a a) = 0$$

for any $a \leq a_b$ and therefore

$$\lambda_a = -\alpha(n)\delta a\tilde{g}(a;n) - \mu_a a \tag{136}$$

for any $a \leq a_b$. In particular, we have

$$\lambda_{a_0} = -\alpha(n)\delta a_0\tilde{g}(a_0;n) - \mu_{a_0}a_0. \tag{137}$$

If $a_0 = 0$, then the above implies that $\lambda_{a_0} = 0$. Next, applying Lemma 6 to the special case $a = a_0$, we have

$$\lambda_{a_0} = \alpha(n)(1-\delta) + \Sigma_{a_c} + \int_{a_0}^{\bar{a}} \eta_x dx. \tag{138}$$

Therefore, if $a_0 = 0$, we have $\lambda_{a_0} = \alpha(n)(1-\delta) + \Sigma_{a_c} + \int_{a_0}^{\bar{a}} \eta_x dx = 0$. ■

Proof of Lemma 9

For all $a \leq a_b$, we have $q_a = 0$ and $q'(a) = 0$. For all a greater than or equal to a_c, q_a is constant and thus $q'(a) = 0$. For $a \in (a_b, a_c)$, implicit differentiation of

$$(a - \phi(a))u'(q_a) = c'(q_a) \tag{139}$$

yields

$$q'(a) = \frac{-[1-\phi'(a)]u'(q_a)}{[a-\phi(a)]u''(q_a) - c''(q_a)} \tag{140}$$

where $\phi(a)$ can be simplified to:

$$\phi(a) = -\left(\frac{(1-\delta)(1-\tilde{G}(a;n)) + \frac{i}{\alpha(n)}}{\delta\tilde{g}(a;n)}\right). \tag{141}$$

Differentiating the above yields

$$\phi'(a) = \frac{1-\delta}{\delta} + \frac{\left[(1-\delta)(1-\tilde{G}(a;n)) + \frac{i}{\alpha(n)}\right]\tilde{g}'(a;n)}{\delta\tilde{g}(a;n)^2}. \tag{142}$$

Since $u'(q_a) > 0$ and $u''(q_a) < 0$ and $c''(q_a) > 0$ and $a - \phi(a) > 0$, we have $q'(a) \geq 0$ provided that $\phi'(a) < 1$. Rearranging, this is true provided that

$$\left(\frac{(1-\delta)(1-\tilde{G}(a;n)) + \frac{i}{\alpha(n)}}{\tilde{G}(a;n)}\right)\left(\frac{\tilde{g}'(a;n)\tilde{G}(a;n)}{\tilde{g}(a;n)^2}\right) < 2\delta - 1. \tag{143}$$

To prove this, we first show that

(144) $$\frac{(1-\delta)(1-\tilde{G}(a;n))+\frac{i}{\alpha(n)}}{\tilde{G}(a;n)} < 2\delta - 1.$$

Rearranging the above and simplifying, this is equivalent to

(145) $$\delta(1+\tilde{G}(a;n)) > 1 + \frac{i}{\alpha(n)}.$$

For any $a \in (a_b, a_c)$, this is true if $\delta \geq 1 + \frac{i}{\alpha(n)}$, which is true since

(146) $$\delta = \frac{1-\tilde{G}(a_b;n)+\frac{i}{\alpha(n)}}{1-\tilde{G}(a_b;n)-a_b\tilde{g}(a_b;n)} \geq 1 + \frac{i}{\alpha(n)(1-\tilde{G}(a_b;n))}.$$

Next, we prove that $G''(a) \leq 0$ is a sufficient (but not necessary) condition for

(147) $$\frac{\tilde{G}(a;n)\tilde{g}'(a;n)}{\tilde{g}(a;n)^2} \leq 1.$$

Combining (55) with (4) yields

(148) $$\frac{\tilde{G}(a;n)}{\tilde{g}(a;n)} = \frac{1-e^{-nG(a)}}{ng(a)}$$

and also, differentiating (55), we obtain

(149) $$\frac{\tilde{g}'(a;n)}{\tilde{g}(a;n)} = \frac{g'(a)+ng(a)^2}{g(a)}.$$

Combining the above two expressions, we have

(150) $$\frac{\tilde{G}(a;n)\tilde{g}'(a;n)}{\tilde{g}(a;n)^2} = \left(1-e^{-nG(a)}\right)\left(\frac{g'(a)}{ng(a)^2}+1\right).$$

Since $1-e^{-nG(a)} \leq 1$, it suffices to show that $\frac{g'(a)}{ng(a)^2} \leq 0$. This is true provided that $g'(a) \leq 0$. Therefore, if $G''(a) \leq 0$ then $q'(a) > 0$ for all $a \in (a_b, a_c)$. ■

10.4 Other proofs for Section 6

Proof of Lemma 3

Part 1. Follows from the facts that $v_a \geq 0$ for all a and $\dot{v}_a = u(q_a) \geq 0$.

Part 2. For the first part, let $f(a) = \frac{z}{\gamma} - au(q_a) + v_a$. Constraint (72) binds if and only if $f(a) = 0$. Differentiating, we have $f'(a) = -(u(q_a) + au'(q_a)q'(a)) + \dot{v}_a$. Using $\dot{v}_a = u(q_a)$, this implies that $f'(a) = -au'(q_a)q'(a)$. Since $u'(q_a) > 0$ and $q'(a) \geq 0$ is a constraint, we have $f'(a) \leq 0$. Therefore, there exists a unique $a_c \in [a_0, \bar{a}]$ such that $f(a) = 0$ and constraint (72) binds for all $a \in [a_c, \bar{a}]$.

For the second part, if $a \in [a_c, \bar{a}]$ then $\frac{z}{\gamma} = au(q_a) - v_a$. Differentiating, we have $au'(q_a)q'(a) = 0$ for all $a \in [a_c, \bar{a}]$. Since $u'(q_a) > 0$ and $q'(a) \geq 0$ is a constraint, this requires $q'(a) = 0$ and thus q_a is constant on $[a_c, \bar{a}]$.

Part 3. It is clear that $a_0 \leq a_b$ and $a_c \leq \bar{a}$. It remains only to show that $a_b \leq a_c$. We have $q_{a_b} = 0$ while $q_{a_c} > 0$, so $q_{a_b} \leq q_{a_c}$ and thus $a_b \leq a_c$ because $q'(a) \geq 0$. ■

Proof of Lemma 4

In the limit as $n \to 0$, we have $\tilde{G}(a;n) \to G(a)$ by Lemma 1. As $n \to 0$, we have $i/\alpha(n) \to \infty$ so $\delta \to \infty$. Using (29), $1/\delta \to 0$ implies that $\phi(a) \to \frac{1-G(a)}{g(a)}$ on $(a_b, a_c]$. We have $a = \phi(a)$ for all $a \leq a_b$, where $\phi(a) = \frac{1-G(a)}{g(a)}$. From Lemma 7, we obtain $a_b = \phi(a_b)$, or $\varepsilon_\pi^0(a_b) = 1$ where $\varepsilon_\pi^0(a) \equiv \frac{ag(a)}{1-G(a)}$. Since we assume that G has a weakly increasing hazard rate, $\frac{g(a)}{1-G(a)}$ is weakly increasing in a and thus $\varepsilon_\pi^0(a)$ is strictly increasing in a. So there is a unique solution $a_b \in [a_0, \bar{a}]$, given that $\varepsilon_\pi^0(a_0) = 0$ and $\lim_{a \to \bar{a}} \varepsilon_\pi(a) = \infty$. Finally, in the limit as $n \to 0$, the condition for a_c reduces to

$$(\bar{a} - a_c)[1 - G(a_c)] = \bar{a}\left[1 - G(a_b) - a_b g(a_b)\right]. \tag{151}$$

Therefore, since $\varepsilon_\pi^0(a_b) = 1$, we have $a_c = \bar{a}$. ■

Proof of Corollary 1

From Proposition 2, for any $a \in [a_b, a_c]$, and thus for any $a \in [a_0, a_c]$ since $a_b = a_0$,

$$(a - \phi(a))u'(q_a) = c'(q_a) \tag{152}$$

where

$$\phi(a) = -\left(\frac{1-\delta}{\delta}\right)\left(\frac{1-\tilde{G}(a;n)}{\tilde{g}(a;n)}\right) - \frac{i}{\alpha(n)\delta\tilde{g}(a;n)}. \tag{153}$$

Using expression (101) for δ, and the fact that $a_b = a_0$ implies $\int_{a_0}^{\bar{a}} \eta_x dx = 0$ since $\eta_a = 0$ for all a, we obtain

$$\delta = 1 + \frac{i}{\alpha(n)}. \tag{154}$$

Substituting into the above expression for $\phi(a)$ and simplifying, we obtain (36). ■

10.5 Proofs for Section 7

We first present a lemma that is used to prove Proposition 3.

Lemma 13. *In any equilibrium where $i > 0$,*

1. *There exists a unique cutoff $a_p \in [a_b, \bar{a}]$ such that (i) if $a_p \leq a_c$, there is underconsumption for all $a \in (a_0, a_p)$ and overconsumption for all $a \in (a_p, a_c)$, and (ii) if $a_p > a_c$, there is underconsumption for all $a \in (a_0, a_p)$.*

2. *There exists a unique cutoff $a_d \in [a_b, \bar{a}]$ such that (i) if $a_c \leq a_d$, there is overconsumption for all $a \in [a_c, a_d)$ and underconsumption for all $a \in (a_d, \bar{a}]$, and (ii) if $a_c > a_d$, there is underconsumption for all $a \in [a_c, \bar{a}]$.*

Proof. *Part 1. (i)* For all $a \in (a_0, a_b]$, there is underconsumption, i.e. $q_a < q_a^*$, since $q_a = 0$ but $q_a^* > 0$. For $a \in (a_b, a_c]$, we have $a - \phi(a) = c'(q_a)/u'(q_a)$ and $a = c'(q_a^*)/u'(q_a^*)$, where $c'(q)/u'(q)$ is increasing in q, so $q_a < q_a^*$ (i.e. underconsumption) for any $a \in (a_b, a_c]$ if and only if $\phi(a) > 0$. Rearranging (29) yields

$$\phi(a) = -\left(\frac{(1-\delta)(1-\tilde{G}(a;n)) + \frac{i}{\alpha(n)}}{\delta\tilde{g}(a;n)}\right), \tag{155}$$

and therefore $\phi(a) > 0$ if and only if

$$-\left((1-\delta)(1-\tilde{G}(a;n)) + \frac{i}{\alpha(n)}\right) > 0. \tag{156}$$

Rearranging, $\phi(a) > 0$ if and only if

$$\tilde{G}(a;n) < 1 + \frac{i}{\alpha(n)(1-\delta)}. \tag{157}$$

Since $\tilde{G}'(a;n) = \tilde{g}(a;n) \geq 0$, and $\tilde{G}(a_0;n) = 0$ and $\tilde{G}(\bar{a};n) = 1$, while $1 + \frac{i}{\alpha(n)(1-\delta)} \in [0,1]$, there exists a unique cut-off $a_p \in (a_b, \bar{a}]$ such that $\phi(a) > 0$ and there is underconsumption for all $a \in (a_0, a_p)$ where a_p satisfies

$$\delta = 1 + \frac{i}{\alpha(n)[1 - \tilde{G}(a_p;n)]} \tag{158}$$

provided that $a_p \leq a_c$. If $a \in (a_p, a_c)$ then $\phi(a) < 0$ and there is overconsumption. *(ii)* If $a_p > a_c$, the range of overconsumption (a_p, a_c) is empty and we have underconsumption for all $a \in (a_0, a_p)$.

Part 2. *(i)* For all $a \in [a_c, \bar{a}]$, $q_a = q_{a_c}$ where $a_c - \phi(a_c) = c'(q_{a_c})/u'(q_{a_c})$ and $a = c'(q_a^*)/u'(q_a^*)$. Since $c'(q)/u'(q)$ is increasing in q, we have $q_a > q_a^*$ (i.e. overconsumption) if and only if $a < a_c - \phi(a_c)$. Defining $a_d \equiv a_c - \phi(a_c)$, we have overconsumption for $a \in [a_c, a_d)$ and underconsumption for $a \in (a_d, \bar{a}]$. *(ii)* If $a_d < a_c$, the interval $[a_c, a_d)$ is empty and we have underconsumption for all $a \in [a_c, \bar{a}]$. ■

Proof of Proposition 3

Part 1. Suppose that $a_d = \max\{a_c, a_d\}$. Follows from combining Parts 1 and 2 of Lemma 13 if $a_p \leq a_c \leq a_d$. Suppose that $a_c = \max\{a_c, a_d\}$. Follows from combining Parts 1 and 2 of Lemma 13 if $a_p \leq a_c$ and $a_d < a_c$.

Part 2. If $a_p > a_c$, then $\phi(a) > 0$ for all $a \in (a_0, a_p)$ from Part 1 *(ii)* in Lemma 13. In particular, $\phi(a_c) > 0$, so we get $a_c > a_d$. The rest follows from combining Parts 1 and 2 in Lemma 13. If $a_p = a_c$, the result follows from Part 1.

Part 3. If $a_b = a_0$ then $\delta = 1 + \frac{i}{\alpha(n)}$ and (158) implies $\tilde{G}(a_p;n) = 0$ and thus $a_p = a_b = a_0$. Since $a_p \leq a_c$, Part 1 implies there is overconsumption on (a_0, a_u) and underconsumption on $(a_u, \bar{a}]$ where $a_u = \max\{a_c, a_d\}$. Since $\phi(a_c) < 0$ by (36), we have overconsumption at a_c. Therefore, $a_c < a_u$ and $a_u = a_d$. ■

Proof of Corollary 2

Part 1. Follows from Part 1 of Proposition 2.

Part 2. Starting with equation (31) in Proposition 2, setting $i = 0$ implies $a_c = \bar{a}$.

Part 3. Setting $i = 0$ in expression (30) for δ in Proposition 2, we obtain

$$\delta = \frac{1}{1 - \varepsilon_\pi(a_b; n)}. \tag{159}$$

Setting $i = 0$ in expression (29) for $\phi(a)$ in Proposition 2, and substituting (159) into (29), we obtain (41).

Part 4. Parts 5-8 from Proposition 2 also hold. ■

Proof of Proposition 5

In any full-trade equilibrium where $a_b = a_0$, letting $i \to 0$ gives same allocation as planner. If $a_b = a_0$, then $\varepsilon_\pi(a_b) = \varepsilon_\pi(a_0) = 0$ and Corollary 2 implies that q_a satisfies $au'(q_a) = c'(q_a)$ for all $a \in [a_0, \bar{a}]$, which is equivalent to the planner's FOC (7). Also, we know that $\alpha'(n)\tilde{s}(n) + \alpha(n)\tilde{s}'(n) = k$, which is equivalent to the planner's FOC (8). Finally, buyers always choose the highest quality seller in any meeting and therefore the distribution of chosen goods is equal to the distribution of the maximum, given by (34), which is the same as the distribution of goods chosen by the planner. Therefore, $i \to 0$ gives the same allocation as the planner. Conversely, in order to have $q_a = q_a^*$ for all a, we must have $a_b = a_0 = 0$ (except in the limit as $n \to \infty$). ■

Proof of Proposition 6

For any $a \in (a_0, a_b]$, the efficient quantity is not traded even when $i \to 0$ since $q_a^* > 0$ but $q_a = 0$. The efficient quantity is traded at $a_0 = 0$ since $q_0 = q_0^* = 0$. To get the efficient quantity at $i \to 0$ for any $a \in (a_b, \bar{a}]$, Corollary 2 implies that we require $\frac{\varepsilon_\pi(a_b)}{\varepsilon_\pi(a)} = 0$. This is true for $a \in (a_b, \bar{a}]$ only when $a = \bar{a}$ since $q_{\bar{a}} = q_{\bar{a}}^* = 0$. In general, for any $a \in (a_0, \bar{a})$, $q_a \neq q_a^*$. Since $\frac{\varepsilon_\pi(a_b)}{\varepsilon_\pi(a)} \geq 0$, we have $1 - \frac{\varepsilon_\pi(a_b)}{\varepsilon_\pi(a)} \geq 1$ and there is underconsumption for $a \in (a_0, \bar{a})$. ■

Proof of Proposition 7

At the Friedman rule, entry is efficient if the equilibrium is full trade. In any partial-trade equilibrium, we know from Proposition 6 that $q_a^* > q_a$ for any $a \in (a_0, \bar{a})$

at the Friedman rule. The equilibrium n satisfies

$$\alpha'(n)\tilde{s}(n;\{q_a\}_{a\in[a_0,\bar{a}]}) + \alpha(n)\tilde{s}'(n;\{q_a\}_{a\in[a_0,\bar{a}]}) = k \tag{160}$$

and the efficient n^* satisfies

$$\alpha'(n^*)\tilde{s}(n^*;\{q_a^*\}_{a\in[a_0,\bar{a}]}) + \alpha(n^*)\tilde{s}'(n^*;\{q_a^*\}_{a\in[a_0,\bar{a}]}) = k. \tag{161}$$

We know that $q_a^* > q_a$ for any $a \in (a_0, \bar{a})$, but we cannot infer anything about whether there is under-entry ($n < n^*$), over-entry ($n > n^*$), or efficient entry ($n = n^*$). We can find examples of equilibria in which each of these three possibilities occurs. ■

Proof of Proposition 8

Part 1. Suppose $\phi(a) < 0$, i.e. there is overconsumption at any $a \in (a_b, a_c)$. Using the definition of p_a given by (42), we have

$$p_a = \frac{d_a/\gamma}{q_a} = \frac{au(q_a) - v_a}{q_a}. \tag{162}$$

Since $q'(a) > 0$ on (a_b, a_c) or $q \in (0, q_{a_c})$, there is a function $p : (0, q_{a_c}) \to \mathbb{R}_+$ and $p'(q) = \frac{dp}{d_a}\frac{d_a}{d_q}$. Differentiating the above yields

$$p'(q) = \frac{1}{q'(a)}\left(\frac{au'(q_a)q'(a) + u(q_a) - v'(a)}{q_a} - \frac{q'(a)[au(q_a) - v_a]}{q_a^2}\right). \tag{163}$$

Simplifying, and using the fact that $v'(a) = u(q_a)$, we obtain

$$p'(q) = \frac{au'(q_a)q_a - d_a/\gamma}{q_a^2}. \tag{164}$$

We have $p'(q) < 0$ if and only if $au'(q)q < d_a/\gamma$. Using the fact that $u'(q) = c'(q)/[a - \phi(a)]$, this is true provided that

$$\frac{c'(q)q}{d_a/\gamma} < \frac{a - \phi(a)}{a}. \tag{165}$$

We know that $d_a/\gamma \geq c(q_a)$, so it suffices to show that

$$\frac{c'(q)q}{c(q)} < \frac{a-\phi(a)}{a}. \tag{166}$$

Since we assume $\frac{c'(q)q}{c(q)} \leq 1$, $\frac{a-\phi(a)}{a} > 1$ suffices, which is true if $\phi(a) < 0$.

Part 2. It follows from Proposition 3 that any underconsumption range within the range (a_b, a_c) is either (a_b, a_p) or (a_b, a_c). Thus, for any a, we have $\phi(x) \geq 0$ for *all* $x \in (a_b, a)$. Since $q'(a) > 0$ on (a_b, a_c) or $q \in (0, q_{a_c})$, there is a function $p : (0, q_{a_c}) \to \mathbb{R}_+$ and, using definition (42), and $v_a = \int_{a_0}^{a} u(q_x)dx$, yields

$$p(q) = \frac{a(q)u(q) - \int_{a_0}^{a(q)} u(q)dx}{q}. \tag{167}$$

Starting with $[a - \phi(a)]u'(q_a) = c'(q_a)$ and then integrating with respect to q_a on $[0, q_a]$, and using integration by parts, yields

$$p(q) = \frac{\phi(a(q))u(q) + c(q) + F(a(q))}{q} \tag{168}$$

where $F(a) \equiv \int_{a_b}^{a} \phi(x)u(q_x)dx$. Differentiating (168) and rearranging yields

$$p'(q) = \frac{[\phi'(a) + \phi(a)]\frac{u(q_a)q_a}{q'(a)} + \left[\frac{c'(q_a)q_a}{c(q_a)} - 1\right]c(q_a) - F(a) + \phi(a)\left[\frac{u'(q_a)q_a}{u(q_a)} - 1\right]u(q_a)}{q_a^2}. \tag{169}$$

By assumption, $\frac{c'(q_a)q_a}{c(q_a)} \leq 1$, so $\left[\frac{c'(q_a)q_a}{c(q_a)} - 1\right]c(q_a) \leq 0$. Also, we assume $\frac{u'(q_a)q_a}{u(q_a)} < 1$ and $\phi(a) > 0$ so we have $\phi(a)\left[\frac{u'(q_a)q_a}{u(q_a)} - 1\right]u(q_a) \leq 0$. Next, $F(a) \geq 0$ since $\phi(x) \geq 0$ for *all* $x \in (a_b, a)$. Therefore, if we consider the first term in (169), since $q'(a) \geq 0$ it is sufficient to show that $\phi'(a) + \phi(a) \leq 0$. To establish $\phi'(a) + \phi(a) \leq 0$ on $a \in (a_b, a_c)$, we use (141) for $\phi(a)$ and (142) for $\phi'(a)$ to write

$$\phi'(a) + \phi(a) = \frac{1-\delta}{\delta} + \phi(a)\left(1 - \frac{\tilde{g}'(a;n)}{\tilde{g}(a;n)}\right). \tag{170}$$

Since $\delta \geq 1$, we have $\frac{1-\delta}{\delta} \leq 0$ and therefore $\phi'(a) + \phi(a) \leq 0$ provided that $\frac{\tilde{g}'(a;n)}{\tilde{g}(a;n)} \geq 1$, since $\phi(a) \geq 0$. Using expression (149), we have $\frac{\tilde{g}'(a;n)}{\tilde{g}(a;n)} \geq 1$, and therefore $p'(q) \leq 0$, provided that condition (43) holds. ■