Abstract
In order to capture the demand externalities associated with taste for variety and firm entry, we integrate monopolistic competition and preferences for variety in the decentralized market of a Lagos-Wright setting, with an ex ante division of buyers and sellers. Matching is multilateral in that each buyer accesses a range of sellers who each supply a unique variety of the good. We investigate the efficiency properties of equilibrium, firm size, and welfare cost of inflation in a pure monetary economy. Markups arise from the taste for variety and create a rent-sharing externality which amplifies the cost of holding money induced by inflation. But with scale economies markups also help pay fixed costs of entry and thereby enable more variety. In the basic model without entry, consumption and firm size is too low and the Friedman rule is the optimal policy, though it implements the first best only as the taste for variety (and markups) approach zero. I extend the model in three important ways: free entry of sellers, free entry and variable markups, and variable search intensity. Equilibrium exists for each model under mild conditions and is generally unique except for the CES model with free entry. There, uniqueness fails to hold if taste for a variety is too high. In that case, with sufficiently low entry costs, there is a stable high-variety equilibrium and an unstable low-variety equilibrium, where the former Pareto dominates the latter. With CES and free entry, the optimal scale of a firm lies between the efficient scale and the equilibrium level depending on the elasticity of the matching function. The Dixit Stiglitz result that the optimal measure of sellers exceeds the equilibrium measure holds with sufficiently low search frictions, high nominal interest rates, and high elasticity of substitution, but does not hold generally. I compute the welfare costs of inflation in the former case using a compensated measure and use parameter values obtained by fitting the theoretical money demand to its empirical counterpart. For markups of 30%, the estimates are about 7.24% without entry and 9.25% with entry. Under variable markups introduced via additively separable preferences, markups decrease with interest rates and attenuate the welfare costs of inflation.

Keywords: money, inflation, search, matching, monopolistic competition, taste for variety, free entry

1. Introduction
Taste for variety can be formalized in many ways but is most naturally conceived in terms of the convexity of indifference curves. This innovation, which follows from a basic property of consumer theory, was the novel feature of Dixit and Stiglitz (1977), and brought about a monopolistic competition revolution in much of macroeconomics. Some of the most interesting implications are the product of the general equilibrium implications of an otherwise simple microeconomic idea. One especially interesting implication is the strategic complementarity between variety of goods and firm entry: if sellers which offer specialized products enter the market, this boosts demand for existing products because of complementarity. This higher demand, in
turn, makes the entry of firms more profitable. Thus, variety and free entry together bring about demand externalities and hence creates a multiplier effect. Moving one level further, taste for variety is an equilibrium object—as opposed to an intrinsic property of preferences—whenever the elasticity of the marginal rate of substitution (and demand) varies with quantity. This equilibrium level of taste of variety plays an important (unique in the case of DS) role in (1) the determination of markups and (2) explaining the extent to which goods are imperfect substitutes. Furthermore, monopolistic competition is consistent with firms posting prices to an anticipated mass of consumers, which reasonably describes a very high fraction of trade. Given that monetary theory is extensively concerned with equilibrium trade, firm entry, market power, and price setting, it stands to reason that taste for variety and monopolistically competitive firms selling differentiated products should be an important part of the story. Yet these ingredients do not feature prominently in new monetarist models.

A major reason is that monetary theory is primarily concerned with carefully describing the social role of money in terms of essential frictions. Specifically, in the presence of limited commitment and anonymity, money is essential because it overcomes a double coincidence of wants problem (Kocherlakota 2005). For the most part, the new monetarist framework follows the approach of Diamond (1982) and describes an economy with pairwise meetings and search frictions. Prices are typically determined through bargaining, though alternative approaches have been studied. In particular, there is competitive search (Moen 1997, Mortensen and Wright 2002, Faig and Jerez 2006, and Rocheteau and Wright 2005, 2009) and auctions (Galenenanos and Kircher 2008). Furthermore, Rocheteau and Wright (2005), henceforth, RW; and Laing, Li, and Wang (2007), hereafter LLW, show that frictions which render money essential are compatible with competitive pricing and monopolistic competition, respectively.

The key innovation of this paper, which is partially anticipated by LLW, is the integration of the frictions which make money useful in equilibrium alongside search, taste for variety and specialization, and free entry. Said otherwise, we integrate the insights of DS—and the concomitant demand externalities—with a deep model of money. In so doing, we pave the way for a more extensive integration of monetary theory into other parts of macroeconomics that have heavily built on Dixit-Stiglitz monopolistic competition: international trade, new Keynesian theory, growth theory and economic geography.¹

We illustrate this framework by using it to reexamine the welfare costs of inflation, building on the seminal work of Lucas (2000), as well as how monetary policy affects number of varieties (and firm entry), firm size, and efficiency properties of equilibrium. We find that the Friedman rule is optimal both with and without entry for CES and that the welfare costs of inflation associated with 30% markups are about are about 7.24% without entry and 9.25% with entry. This is among the highest obtained in the literature. For instance, in Rocheteau Wright 2009, the estimated welfare cost is 7.41% if \( \theta = 0.2 \) and 3.10% if \( \theta = 0.5 \).

The paper is organized as follows. Section 2 reviews the related literature. Section 3 presents the basic environment, and Section 4 analyzes the equilibrium and social optimum. Section 5 follows DS and allows for free entry of producers with an entry cost. In contrast to the findings of Shi (1997) and Rocheteau Wright (2005), the Friedman rule maximizes equilibrium welfare. The equilibrium measure of producers satisfies the condition that expected profit equals the entry cost. Section 6 examines equilibrium under additively separable preferences, which eliminate complementarity and give rise to variable markups. Section

¹Classic references include Krugman (1979), Krugman (1991), Woodford (2003), Romer (1990) and many others.
2. Related literature

This paper builds on a rich body of work on monetary theory and the model of monopolistic competition developed by Dixit and Stiglitz (1977). A major benchmark for monetary theory is Lagos and Wright (2005), hereafter LW. LW develop a tractable model of divisible money in which each period is divided into two subperiods, one with a decentralized market (DM) and one with a centralized market (CM). Quasilinear utility in the CM pins down the quantity of the CM good and eliminates wealth effects. Thus, there is a degenerate distribution of real balances across agents. LW show that efficiency under generalized Nash bargaining requires the Friedman rule (setting the nominal interest rate to zero) and setting the bargaining power of buyers to one (buyers take all). The first condition arises because the Friedman rule leads to a zero cost of holding real balances, and the second arises because of a holdup problem. The holdup problem is that buyers bring assets from the CM to the DM, and they only invest optimally if they can appropriate the full return. In this setting, with buyers obtaining only a portion of the surplus, the first best is generally not implementable unless instruments besides interest rates are available.

A major advantage of the LW environment is that it is compatible with a variety of market structures. Hence, the LW formalism is useful for constructing a taxonomy of market structures, allowing for deeper integration into the rest of macroeconomics. One can incorporate different types of bargaining, as proportional bargaining (Aruoba, Rocheteau, and Waller 2007); mechanism design (Hu, Kennan, and Wallace 2009) for the purposes of normative analysis; competitive search (Rocheteau and Wright 2005, 2009); auctions; and Walrasian price taking.

RW examine bargaining, Walrasian price taking, and competitive search in detail. In each of these models, they expand upon LW by allowing entry of sellers and dividing buyers and sellers ex ante, which serves as a simple way to incorporate the extensive margin. They compare equilibrium to the social optimum under these different market structures. Under bargaining, the quantity traded and entry are inefficient. Inflation implies a first-order welfare loss. Under competitive equilibrium, the Friedman rule implies efficiency along the intensive margin but not the extensive margin. Efficiency would only hold under the extensive margin under a Hosios-like condition in which the congestion externality balances out the thick market externality. In competitive search, the Friedman rule achieves the first best: the welfare loss due to inflation, however, is of second order.

The case of Walrasian price taking in RW, in which search frictions arise from a queue of buyers and sellers, is especially relevant for our purposes. As I later show, the model I develop is an extension of RW

---

2Furthermore, in a separate working paper, available upon request, I endogenize the number of varieties per firm in a way similar to Dong (2010). Firms specializing in one variety just keeps the entry margin as easy as possible, but the extension to measures of varieties per firm is straightforward and does not affect the results of the current paper substantively.

3More precisely, the surplus of the buyer is not a monotonic function of the total surplus for $\theta < 1$. 
with Walrasian price taking. The key differences are both the relaxation of price taking behavior and the incorporation of search frictions along consumption variety.

Of the two major monetary models with monopolistic competition, neither has integrated monopolistic competition and search frictions in the DM of a LW-type model. Laing, Li, and Wang (2007), henceforth LLW, use multilateral matching between a continuum of buyers and sellers. Search frictions arise in terms of limited consumption variety rather than probability of participation. LLW do not consider idiosyncratic uncertainty and hence focus on the monopolistic goods sector. Households provide labor competitively to firms. LLW take CES preferences over goods and leisure where labor and search effort satisfy a time resource constraint. As we do here, they consider fixed search effort and variable search effort. However, they do not consider the role of congestion in search: higher search by other buyers does not have a negative effect on the measure of sellers a given buyer can contact. They have two primary findings in the case where the matching depends on search. First, if there is sufficient complementarity between goods and labor, there is a unique steady state equilibrium where inflation reduces labor, search effort, and output. If labor and goods are highly substitutable, then there is a unique steady state in which inflation increases labor, search effort, and output if there is sufficient taste for variety and decreases them otherwise.

The other major approach is that of Aruoba and Schorfheide (2009). They use monopolistic competition in the CM rather than the DM. They modify the CM as a variant of the new Keynesian model, including monopolistic competition and nominal rigidities. Their primary problem is to integrate the welfare costs of inflation from in terms of its productive activity (the Friedman channel) and relative price distortions (the new Keynesian channel). The project represents a bold effort at integration. One problem is that in keeping search frictions in the DM and nominal rigidities in the CM, they do not fully integrate the effects of monopolistic competition with monetary search frictions. For instance, the setup does not allow money growth to influence the variety of monopolistic goods purchased.

Dong (2010) models inflation and variety in an LW-type environment in which buyers receive i.i.d preference shock each period for a specific type of variety and sellers produce a unique set of special goods, which is increasing in investment in the CM. Product variety is endogenous through firm investment instead of firm entry. In effect, buyers have a stochastic taste for variety across periods rather than a deterministic taste for variety within a period, when goods are bought and sold. Dong considers both Nash bargaining and directed search with price posting as trading mechanisms. In both cases, inflation reduces both quantity and variety. The Friedman rule is the best policy in both mechanisms and attains the first best for price posting. This means that Dong’s model cannot generate price posting and markups simultaneously. Furthermore, the welfare costs of inflation due specifically to loss in variety are substantial for Nash bargaining but negligible for price posting. The model I present features price posting with markups from goods that results directly from the taste for variety.

The benchmark model of monopolistic competition is by Dixit and Stiglitz (1977). Contrary to earlier models, they formalized buyers having a taste for variety in terms of the convexity of indifference curves. The core model uses CES preferences, which has become an integral part of many branches of macroeconomics due to its tractability. Under CES, they find that firm quantity is optimal but that there are too few firms (too few varieties). They also considered a form of variable markups and showed that if optimal firm production is less than the equilibrium production, then the optimal number of firms exceeds the equilibrium number. Depending on preferences, the optimum may have more firms and larger firms. Variable markups were used by Krugman (1979) in additively separable form to analyze growth, trade, and factor mobility. Crucial to his
analysis is that the price elasticity of demand increases with price. Zhelobodko, Kokovin, Parenti, and Thisse have analyzed additively separable utility in depth to provide a complete market characterization with free entry. Two other forms of variable markups are quasilinear quadratic preferences (Ottaviano, Tabuchi, and Thissé 2002), and translog preferences (Feenstra 2003). Melitz (2003) uses quasilinear quadratic preferences to analyze trade flows under heterogeneous firms. Additively separable preferences, translog preferences, and quasilinear quadratic preferences can be generalized to a class of preferences considered by Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2012). Bilbiie, Ghironi, and Melitz (2007) use general homothetic preferences, which nest CES and translog to study business cycle dynamics with product variety and entry. These class of preferences can be summarized in terms of quantity and price indices. I am not aware of any study of the interplay of monetary theory and monopolistic competition with variable elasticity of demand to date.

3. Basic environment

The environment builds on the new monetarist model by Rocheteau and Wright (2005) and Lagos and Rocheteau (2005). Time is discrete and the horizon is infinite. Each period has two subperiods: in the first subperiod there is a decentralized market (DM) and in the second subperiod there is a centralized market (CM). In the decentralized market a continuum of goods (DM goods) indexed on $[0,1]$ are consumed and produced. In the CM a general good is consumed and produced using labor hours $h$, where $h$ produces $x$ units of the CM good. The period utility function of the buyer is given by

$$U^b(x, h, q(j)) = \psi u \left[ \left( \int_{j \in [0,1]} q(j)^{\eta-1} \, dj \right)^{\eta\eta^{-1}} \right] + U(x) - h$$

where $u(0) = 0, u'(0) = \infty, u'(q) > 0$, and $u''(q) < 0$ for $q > 0$. $U(\cdot)$ satisfies the same conditions as $u(\cdot)$ and $\psi$ is a preference shock. I consider $\text{Prob}(\psi = 1) = \sigma$ and $\text{Prob}(\psi = 0) = 1 - \sigma$. Moreover, $\eta \in (1, \infty)$. These preferences imply constant elasticity of substitution of the individual DM goods. Notice that the CM good is analogous to the outside sector in DS.

The period utility function of the seller is

$$U^s(x, q) = -c(q) + U(c) - h$$

where $c(0) = c'(0) = 0, c'(q) > 0, c''(q) \geq 0$ for $q > 0$.

Matching is multilateral: each buyer matches with a measure of sellers, and each seller produces goods for a measure of buyers. There is a mass of measure 1 of buyers and a mass of measure $\mu$ of sellers. Each buyer matches with a set of measure $\alpha(\kappa) \leq \mu$ of sellers (and varieties), where $\kappa = \mu/\sigma$. The measure of buyers serviced by each firm is given by $\alpha(\kappa)/\kappa$. The probability that a particular seller is chosen by a particular

---

4 It is immediate that the Dixit-Stiglitz aggregator, the quantity in brackets, is concave with respect to $q(j)$. The restriction $\eta > 1$ is necessary for demand to be elastic (so that marginal revenue of firms is positive). As $\eta \to \infty$, the DM goods become perfect substitutes with each other.

5 One restriction between this setup and DS is that here there are quasilinear preferences between the monopolistic competitive sector and the outside sector, whereas in DS there are general homothetic preferences. The restriction, as in LW, serves to make distributions of real balances degenerate across agents.
buyer is given by $\alpha(\kappa)/\mu$. I require that $\lim_{\mu \to 0} \alpha(\kappa)/\mu = 1$. By L'Hopital's rule, this requires $\alpha'(0) = 1$. In this section we fix $\mu$ (and hence $\kappa$). In Section 5, we allow the measure of sellers $\mu$ to vary. Note that search frictions depend on an idiosyncratic component through $\sigma$ and a non-idiosyncratic component through $\alpha$. $\alpha = \sigma = 1$ leads to Dixit-Stiglitz with CES preferences as a special case. $\sigma < 1$ and $\eta = \infty$ leads to Rocheteau Wright with competitive pricing and a fixed set of sellers.

Buyers and sellers differ in their preferences and production possibilities. During the CM both have the ability to produce and wish to consume. In the DM, buyers want to consume but cannot produce whereas sellers are able to produce but do not wish to consume.  

I assume agents are anonymous and that there are no forms of commitment or public memory that would render money inessential. Fiat money is costless to produce, intrinsically useless, perfectly divisible, and storable. The ex ante division between buyers and sellers together with anonymity rule out double coincidence of wants, generating an essential role for money. The gross growth rate of the money supply is constant over time and equal to $\gamma$: $M_{t+1} = \gamma M_t$. New money is injected (or withdrawn if $\gamma < 1$) by lump-sum transfers (or taxes). These transfers take place during the CM and without loss of generality they go only to buyers.

4. Equilibrium and the social optimum

It is useful to write utility of the consumer as a function of aggregate consumption in the DM (which depends on trading frictions) and the CM. This is possible because constant elasticity of substitution among DM goods means that the consumer is indifferent to any subset of goods with the same measure. With no loss of generality, I relabel the index of the $\alpha$ measure of goods consumed by a buyer on $[0, \alpha]$.

$$\bar{q} = \left[ \int_0^\alpha q(j) \frac{\eta - 1}{\eta} dj \right]^{\frac{\eta}{\eta - 1}}$$

This definition is the integral form of the quantity index in DS with the upper limit of integration reflecting the search friction. The marginal utility of consuming greater variety is given by $u'(\bar{q}) \frac{\eta - 1}{\eta} = \frac{\eta - 1}{\eta} u'(q(\alpha)) (\bar{q} \eta - 1)^{\frac{\eta - 1}{\eta}}$, which approaches $u'(\bar{q})q(\alpha)$ as $\eta \to \infty$ and infinity as $\eta \to 1$. In the former case, there is a pure quantity effect.

I can rewrite the preferences of the buyer as

$$U^b(\bar{q}, x) = u(\bar{q}) + U(x) - h$$

Consider first the problem of a buyer holding $z$ balances when entering the centralized market. A lump-sum transfer equal to $T = \phi_t(M_{t+1} - M_t)$ is given to buyers, where $\phi_t$ is the value of money. I focus on steady state equilibria where aggregate real balances are constant: $\phi_t M_t = \phi_{t+1} M_{t+1}$. In order to hold $z'$ next period, the buyer must accumulate $\gamma z'$, where $\gamma$ is the gross inflation rate $\frac{\phi_{t+1}}{\phi_t}$. The consumer allocates real balances and transfers into spending on the general good and savings for the following DM. Hence, the value function takes $W(z)$ in the CM satisfies

$$W(z) = \max_{z', x, h \geq 0} \left[ U(x) - h + \beta V(z') \right] \quad \text{s.t.} \quad x + \gamma z' = h + z + T \quad (1)$$

---

6This version thus differs from Lagos and Wright (2005) in that buyers and sellers are ex ante different.

7Quasilinear utility in the CM implies no wealth effects from the lump-sum transfer and thereby makes the allocation of transfers immaterial.
Substituting the constraint we can rewrite the value function as

\[ W(z) = z + T + \max_{x \geq 0} \{ U(x) - x \} + \max_{z' \geq 0} \{ \beta V(z') - \gamma z' \} \] (2)

From quasilinear preferences the value function is linear in \( z \) and that the choice of real balances \( z' \) is independent of \( z \). The first order condition with respect to \( h \) after substituting the constraint for \( x \) yields \( U''(x^*) = 1 \). Note that \( T = z(\gamma - 1) \) and from the budget constraint \( h^* = x^* \).

The value function in the DM \( V(z) \) can be expressed as

\[ V(z) = \max_{q(j)} \left\{ \sigma u(\eta q) + \sigma W \left( z - \int_0^\alpha p(j)q(j) dj \right) \right\} \text{ s.t. } \int_0^\alpha p(j)q(j) dj \leq z \] (3)

where the resource constraint reflects lack of credit. If \( \frac{\phi_{x+1}}{\phi_1} < \frac{1}{\beta} \) (\( \gamma > \beta \)), then money is costly to hold and the constraint \( \int_0^\alpha p(j)q(j) dj \leq z \) binds. The value function for the DM can be written as The objective function of the buyers boils down to

\[ \max_{q(j)} \{ \sigma[u(\eta q) - z] - iz \} \] (4)

where \( 1 + i = \frac{\gamma}{\beta} \). The interpretation is that buyers face an opportunity cost of \( i \) on real balances brought in DM regardless of whether the face a preference shock or not. The first order condition yields

\[ \left( 1 + \frac{i}{\sigma} \right) p(j) = u'(\eta q) \left[ \frac{\eta}{q(j)} \right]^{1/\eta} \] (5)

Thus the buyer equates \( (1 + i/\sigma)p(j) \), the cost of acquiring the good taking into account the monetary wedge due to inflation, with the marginal benefit of the good, which is higher with more consumption of the other DM varieties.

Note from (5) we have the relationship

\[ \frac{q(j)}{q(i)} = \left[ \frac{p(i)}{p(j)} \right]^{\eta} \]

where \( \eta \) represents the elasticity of substitution and elasticity of demand. Following Mrazova and Neary (2013), hereafter MN, we define the curvature of demand of variety \( j \) as \( \xi(q_j) = -\frac{p''(q_j)q_j}{p'(q_j)} \). It is easy to directly show that \( \xi(q_j) = \frac{\eta+1}{\eta} \). As MN discuss, elasticity and curvature are sufficient statistics for many comparative statics, a point we will revisit.

Let us now turn to the maximization problem of the monopolistic competitor in the DM. Each firm \( j \) produces a unique type \( j \) and quantity \( q(j) \) for a given consumer. The measure of consumers which match with the firm is \( \alpha(\kappa)/\kappa \). Hence, each firm produces \( \frac{\alpha(\kappa)}{\kappa} q(j) \) overall, which we denote \( q_s(j) \).

\[ \max_{q(j), p(j)} \{ p(j)q_s - c[q_s(j)] \} \]

subject to the inverse demand function given by (5) and given \( \eta \). The solution is given by

\[ p(j) = \frac{\eta}{\eta - 1} c'[q_s(j)] \] (6)

There is a constant markup of price to marginal cost that depends negatively on the elasticity of substitution (positively on product differentiation).\(^8\) Perfect competition is the limiting case as \( \eta \to \infty \). From (5) the

\(^8\)It is well known that, if feasible, nonlinear pricing schemes are more profitable for the firm than linear pricing (i.e. Stiglitz
problem of the firm is strictly concave and thereby admits a unique solution \( q(j) \).\(^9\) That each firm faces the same problem\(^10\), \( q(j) = q \). Furthermore, given identical production, \( \Upsilon = q \alpha(\kappa) \eta/(\eta - 1) \), and \( p(j) = \alpha \) for all \( j \):

\[
p = \frac{u'(\Upsilon)}{1 + i/\sigma} \alpha(\kappa) \eta/(\eta - 1)
\]

I next turn to the social optimum. The formal social planning problem is to choose quantities of variety \( j \) for consumer \( i \) \( q_{i,j} \) to maximize aggregate welfare net of production costs, weighing each individual equally. Let \( \Psi \) denote the set of buyers which has a preference shock (of measure \( \sigma \)), \( A_i \) denote the set of sellers who buyer \( i \) contacts (of measure \( \alpha(\mu) \)), and \( B_j \) denote the set of buyers who approach seller \( j \) (of measure \( \alpha(\kappa)/\kappa \)). The social planning problem is thus given by

\[
\Omega(q_{i,j}) = \max_{q_{i,j}} \left\{ \int_{s \in \Psi} u \left[ \left( \int_{A_i} q_{i,j}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\sigma}{\eta}} \right] di - \int_0^\mu c \left( \int_{B_j} q_{i,j}^\eta di \right) dj \right\}
\]

Given the convexity of the cost function, it is socially optimal for each firm to produce the same quantity \( q_s \). Furthermore, given the concavity of \( \Upsilon \) with respect to \( q(j) \), it is optimal to have \( q(j) = q \) for all \( j \). Given concavity of \( u(\cdot) \) it is optimal to set \( \Upsilon(j) = \Upsilon \) for all \( j \). Noting that aggregate costs are given by \( \mu c(q_s) \), and writing \( \Upsilon = \alpha(\kappa)/\kappa q_s \) the social welfare can be defined simply as a function of \( q_s \):

\[
W(q_s) = \sigma u \left[ \alpha(\kappa) \mu q_s \right] - \mu c(q_s)
\]

The strict concavity of the social welfare function defines a unique solution \( q_s \) given by

\[
\frac{\alpha(\kappa)^{1/(\eta - 1)} u'(\alpha(\kappa)^{1/(\eta - 1)} \mu q_s)}{c'(q_s)} = \frac{\eta}{\eta - 1} \left( 1 + \frac{i}{\sigma} \right)
\]

Equation (10) determines the equilibrium level of \( q_s \).

**Definition 1.** A (steady state) equilibrium is a list \((q_s, p)\) satisfying

\[
\frac{\alpha(\kappa)^{1/(\eta - 1)} u'(\alpha(\kappa)^{1/(\eta - 1)} \mu q_s)}{c'(q_s)} = \frac{\eta}{\eta - 1} \left( 1 + \frac{i}{\sigma} \right)
\]

Proposition 1 (Existence and uniqueness). There is a unique equilibrium.

\(1977\). For instance, a firm would like to set two-part pricing \( p(q) = A + c'(q)q \), where \( A \) is the consumer surplus. This both increases profits and eliminates the inefficiency from monopolistic competition. In practice, however, we observe sizeable markups and few fixed fees. One problem is that this provides consumers an incentive to combine purchases into as few transactions as possible and resell goods. For the problem at hand, linear pricing is a reasonable approximation, but nonlinear pricing can be included in a multisector model to take into account that the feasibility thereof depends strongly on the industry. In this particular setting, however, two-part pricing which fully extracts consumer surplus would drive real balances to and lead to a breakdown of trade.

\(^9\)The second order condition is \( 2p'(q_s) + q_s p''(q_s) - c''(q_s) < 0 \), which simplifies to \( \eta > 1 \) after using \( \xi = \frac{\eta + 1}{\eta} \) and inserting the first order condition.

\(^10\)This is a result of the fact the elasticity of substitution is independent of specific varieties of the DM good and that the costs functions are the same for each firm.
Proof. Define the left hand side of (10) as \( G(z) \). Because \( u'(\cdot) < 0, c'(\cdot) > 0 \), \( G(z) \) is decreasing. Due to the Inada conditions, \( G(0) = \infty \) and \( G(\infty) = 0 \). Hence, there is a unique value \( z^* \) for which \( G(z^*) = \frac{\eta}{\eta - 1} \left( 1 + \frac{1}{\sigma} \right) \). \( q^* \) is given by \( z^* = \alpha \frac{\eta}{\eta - 1} q^* \). Given \( q^* \), \( p^* \), is uniquely defined by (11).

Note that the marginal markup can be decomposed into a rent-sharing externality \( \frac{\eta}{\eta - 1} \) and cost of holding balances \( 1 + \frac{1}{\sigma} \). Thus, inefficiency arises from (1) markups, which induce a rent sharing externality; (2) a positive nominal interest rate, which makes money costly to hold; and (3) \( \sigma < 1 \), which creates idiosyncratic uncertainty that is unresolvable until after acquiring money.

Note that the Friedman rule is the best policy but does not attain the first best. Thus, efficiency is achievable only asymptotically. The case where \( \eta \to \infty \) corresponds to RW. These results are intuitive. Money allows for beneficial trade in the DM and is costless to produce. Hence, its optimal price, the nominal interest rate, should be 0, provided there are no other externalities. Moreover, there is a rent sharing externality that depends positively on the markup and which amplifies the welfare cost of inflation. Since the buyer’s share of the surplus is non-monotonic under monopolistic competition, the rent sharing externality does not vanish as \( i \to 0 \). We summarize this discussion in several propositions.

**Proposition 2.** An increase in \( i \) leads to a decrease in equilibrium values of \( q_s, \bar{q}, q \).

By making money more costly to hold, a higher interest rate decreases \( q_s \), which has a one-to-one relationship with \( q \) and \( \bar{q} \).

Denote by \( q_s^* \) the solution to the maximization of the social welfare function. Also define the output gap as the relative difference between the equilibrium quantity and the socially optimal quantity: \( \frac{q_s^* - q_s}{q_s^*} \).

**Proposition 3.** Equilibrium is inefficient. Social welfare is maximized at \( i = 0 \).

**Proposition 4.** The output gap is decreasing with \( \eta \). As \( \eta \to \infty \), the marginal markup approaches \( 1 + \frac{1}{\sigma} \).

**Corollary 1.** The policy \( i = 0 \) implements social efficiency asymptotically as \( \eta \to \infty \).

Table 1 compares the marginal markup with proportional bargaining and generalized Nash bargaining, with the derivation sketched in Appendix A.2. For completeness, I also include the marginal markup under additively separable preferences a la ZKPT. Here \( \Theta(q) = \frac{\theta u'(q)}{u'(q) + (1 - \theta)c'(q)} \) is buyer’s share of surplus under Nash and satisfies \( \theta'(q) < 0 \). Furthermore, \( r_u(q) \) is the inverse of the elasticity of demand. Both bargaining mechanisms feature rent sharing externalities. Moreover, both externalities vary with the level of trade and hence the interest rate. However, the rent sharing externality disappears with proportional bargaining as \( i \to 0 \) since the buyer’s surplus is monotonic in the level of trade. Yet inefficiency persists in generalized Nash because consumer surplus is non-monotonic. With CES monopolistic competition, the rent sharing externality is constant because demand elasticity is constant. However, more general preferences allow the demand elasticity to vary and hence the rent sharing externality to vary with the interest rate.
Table 1: Marginal markups

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Marginal markup</th>
<th>Marginal markup: $i = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Nash</td>
<td>$\left{ \frac{\partial'(q) u(q) - c(q) - \partial(q) c'(q)}{c'(q) \left( 1 + \frac{\cdot}{\cdot} \right) } \right} \left( 1 + \frac{\cdot}{\cdot} \right)$</td>
<td>$1 - \frac{\partial'(q) u(q) - c(q)}{\partial(q) c'(q)}$</td>
</tr>
<tr>
<td>Proportional</td>
<td>$\frac{\theta(1 + \frac{\cdot}{\cdot})}{\theta + (1 - \theta) \frac{\cdot}{\cdot}}$</td>
<td>$1$</td>
</tr>
<tr>
<td>CES monopolistic competition</td>
<td>$\frac{\eta}{\eta - 1} \left( 1 + \frac{\cdot}{\cdot} \right)$</td>
<td>$\frac{\eta}{\eta - 1}$</td>
</tr>
<tr>
<td>ZKPT monopolistic competition</td>
<td>$\frac{1}{1 - r_u(q)} \left( 1 + \frac{\cdot}{\cdot} \right)$</td>
<td>$\frac{1}{1 - r_u(q)}$</td>
</tr>
</tbody>
</table>

In fact, as $\eta \to 1$, $q_s \to 0$. As $\eta \to \infty$, $q_s$ satisfies $u'(\kappa q_s) = c'(q_s) \left( 1 + \frac{\cdot}{\cdot} \right)$. It is possible to show $q_s$ increases with $\eta$ for sufficiently high, but the result is not true for all $\eta$.

Comparative statics with respect to parameters $\kappa$ and $\eta$ is more delicate. The following relationship holds between elasticities:

$$
\varepsilon[q_s(\kappa)]\{\varepsilon[u'(\cdot)] - \varepsilon[c'(\cdot)]\} = \frac{ik}{\mu + \kappa i} - \varepsilon[u'(\cdot)] - \frac{1}{\eta - 1} \varepsilon[\alpha(\kappa)] \{1 + \varepsilon[u'(\cdot)]\}
$$

Since $\{\varepsilon[u'(\cdot)] - \varepsilon[c'(\cdot)]\}$ is negative, $\varepsilon[q_s(\kappa)]$ is negative if and only if the right hand side is positive. Since the elasticity of the matching function is at most one, a sufficient condition for the right hand side to be positive is

$$
\frac{ik}{\mu + \kappa i} - \varepsilon[u'(\cdot)] - \frac{1}{\eta - 1} \varepsilon[u'(\cdot)] > 0
$$

In the case of CRRA preferences,(13) reduces to

$$
\frac{i \kappa}{\mu + \kappa} + \frac{\eta \varepsilon - 1}{\eta - 1} > 0
$$

Hence, $\eta \varepsilon \geq 1$ implies $\varepsilon[q_s(\kappa)] < 0$, so that firm quantities decrease with more sellers and increase with a greater percentage of active buyers.

**Lemma 1.** Suppose $u(q) = \frac{q^{1-\varepsilon}}{1-\varepsilon}$. Then $\eta \varepsilon > 1$ implies that $q_s$ decreases with $\kappa$ (an increase in $\mu$ or decrease in $\sigma$). This condition is tight in that for $\eta \varepsilon$ arbitrarily close to 1 from below, there exist matching functions $\alpha(\kappa)$ and values of $\mu, \sigma, i$ for which $q_s$ decreases with respect to $\kappa$.

Equilibrium can be characterized in terms of a price curve and function $\lambda(q_s) = \alpha^{1/(\eta - 1)} u'(\alpha^{\frac{1}{\eta - 1}} \mu q_s)$, which describes marginal utility in terms of the production of each firm.
4.0.1. Consumer surplus

The real balances used for all purchases for one consumer is $z(q_s) = \mu_p c'(q_s)\alpha(\kappa)q = \mu_p \kappa c'(q_s)q_s$ Since $z = \phi M$, the value of money is given by

$$\phi = \frac{\mu_p c'(q_s)\alpha(\kappa)q}{M} \tag{14}$$

The consumer surplus is given by

$$\Omega(q_s) = \sigma u[\alpha(\kappa)\frac{1}{\eta-1}\kappa q_s] - (\sigma + i)z(q_s) \tag{15}$$

Using $\phi'(q_s) = \mu_p \kappa[c''(q_s)q_s + c'(q_s)]$, we find

$$\Omega'(q_s) = \kappa[\sigma u'(\bar{q})\alpha(\kappa)^{\frac{1}{\eta-1}} - (\sigma + i)\mu_p c'(q_s)] - (\sigma + i)\mu_p \kappa c''(q_s)q_s \tag{16}$$

which is proportional to

$$\sigma u'(\bar{q})\alpha(\kappa)^{\frac{1}{\eta-1}} - (\sigma + i)\mu_p c'(q_s) - (\sigma + i)\mu_p \kappa c''(q_s)q_s \tag{17}$$

which is positive for small $q_s$ and negative for sufficiently large $q_s$. Using (10) and rearranging, we find that the change in consumer surplus in equilibrium is proportional to $-(\sigma + i)\mu_p \kappa c''(q_s)q_s < 0$

**Lemma 2.** Consumer surplus is decreasing in equilibrium if and only costs are strictly convex.

Lemma (2) shows that consumer surplus is generally nonmonotonic. The non-monotonicity depends on the convexity of costs, just as in Walrasian price taking. This contrasts with generalized Nash bargaining, in where real balances are a weighted average of the utility of the buyer and costs of the seller, and where the weights are a function of the marginal utility and marginal cost depending on the bargaining power. With Nash bargaining, the buyer’s surplus is falling in equilibrium even if costs are linear. This difference in the type of non-monotonicity does not depend on taste for variety.
5. Entry of sellers

In this section we endogenize the measure of sellers, denoted \( \mu \). Given that there is a unit measure of buyers, \( \mu \) is also the ratio of sellers to buyers. The dependence of \( \alpha \) and \( \alpha_s \) on \( \mu \) reflects search externalities. As \( \mu \) increases, the measure of sellers per shopper increases: there is a thick market externality (buyers can purchase a larger share of DM goods) and a congestion externality (each seller is approached by a smaller measure of buyers). The thick market externality considered here is different from the one in RW where buyers have a greater probability of being matched. Here, there is an expansion of the measure of matches (which occur with probability 1). To emphasize this point and more cleanly compare with DS, I abstract from preference shocks: \( \sigma = 1 \). We compare our results DS, which analyzes the problem of scale versus diversity with free entry of sellers. I examine the same tradeoff but generalized to a monetary economy with search frictions between agents. Hence, \( \kappa = \mu \), and the measure of sellers contacted by each buyer is just \( \alpha(\mu) \). The CES version of DS arises by setting \( \alpha(\mu) = \mu, i = 0, c(\cdot) = c \). The first constraint removes search frictions, the second removes cost of holding real balances, and the third captures scale economies in the simplest way: a fixed cost and constant marginal cost.

Sellers can choose to enter the market in the following DM by paying a cost \( a \) in the current CM. The cost must be paid each CM period for continued entry. The opportunity cost of entry is thus \( k = (1 + \rho)a \), where \( \rho \) is the discount rate \( 1/\beta \).

The measure of the varieties of the DM good a buyer can purchase is given by \( \alpha(\mu) \). The measure of buyers for each seller is \( \alpha_s = \alpha/\mu \), and we assume \( \lim_{\mu \to 0} \alpha_s = 1 \).

5.1. Socially efficient allocations

As in the basic model, firms produce the same quantity \( q_s \) in the social optimum. Using \( \eta = \alpha(\mu)^{1/(\eta - 1)} \mu q_s \), we can write social welfare as a function of \( q_s \) and \( \mu \):

\[
W(q_s, \mu) = u[\alpha^{1/(\eta - 1)} \mu q_s] - \mu [c(q_s) + k]
\]  \hfill (18)

This is the utility of DM consumption of buyers minus the production cost of the sellers and their entry cost. The first order conditions are

\[
[q_s] \quad \alpha^{1/(\eta - 1)} u'(\alpha^{1/(\eta - 1)} \mu q_s) = c'(q_s) \tag{19}
\]

\[
[\mu] \quad \frac{k + c(q_s)}{q_s c'(q_s)} = \left[ \frac{[\epsilon(\alpha(\mu))]}{\eta - 1} + 1 \right] \tag{20}
\]

where \( \epsilon[\alpha(\mu)] \) is the elasticity of the matching function with respect to the measure of sellers. Note that the elasticity is decreasing and bounded above by 1.

This follows easily from the concavity of \( \alpha(\cdot) \). If \( f : \mathbb{R}^+ \to \mathbb{R} \) is increasing, differentiable, and concave, then \( f'(x^*) \leq \frac{f(x^*)}{x^*} \), so that \( x^* f'(x^*) \leq 1 \), or \( \epsilon[f(x)] \leq 1 \). Furthermore, \( x f'(x)/f(x) \) is decreasing with \( x \).

\footnote{A general function that satisfies these conditions is \( \alpha(\mu) = \frac{\mu}{(\mu^{k+1} + 1)^{1/k}} \). The elasticity is given by \( \frac{1}{\mu^{k+1}} \).}

\footnote{This follows easily from the concavity of \( \alpha(\cdot) \). If \( f : \mathbb{R}^+ \to \mathbb{R} \) is increasing, differentiable, and concave, then \( f'(x^*) \leq \frac{f(x^*)}{x^*} \), so that \( x^* f'(x^*) \leq 1 \), or \( \epsilon[f(x)] \leq 1 \). Furthermore, \( x f'(x)/f(x) \) is decreasing with \( x \).}
cost is minimized, because of the value of variety. Here we introduce an important innovation: the optimal scale depends on the effectiveness of forming new matches out of new sellers. In particular, the optimal ratio of average cost to marginal cost is given by the elasticity of the matching function with respect to the measure of sellers divided by $\eta - 1$, which reflects the taste for variety, plus unity. Note that for $\alpha(\mu) = \mu$ the right hand side equals $\frac{\eta}{\eta - 1}$. Also as $\eta \to \infty$, the socially optimal scale becomes the efficient scale.

The ratio of average to marginal cost is an important quantity, so define $\Gamma(q_s) = \frac{k + c(q_s)}{q_s c'(q_s)}$. It is straightforward to show that given $\mu$, the relationship $\Gamma(q_s) = \frac{c[\alpha(\mu)]}{\eta - 1} + 1$ holds for a unique $q_s(\mu)$. $\Gamma(q_s)$ is decreasing everywhere, tends to $\infty$ as $q_s \to 0$, and tends to 0 as $q_s \to \infty$. Note that this holds even if $c(\cdot) = c$. Thus (20) defines an implicit decreasing function $q_s(\mu)$. It is evident that if there is no taste for variety ($\eta \to \infty$) and scale economies, then the optimal measure of firms shrinks to zero, as fixed costs are positive.

5.2. Equilibrium

The problem of the buyers and sellers are identical to the basic case except for the new entry margin. In equilibrium,

$$k = q_s p(j) - c(q_s)$$

(21)

Note that, in contrast to RW, $q_s$ does not depend on the ratio of sellers to buyers, and thus monetary policy cannot affect firm size. This will have important implications for optimal policy, as we shall see. Because sellers face identical problems that admit a unique solution, $q(j) = q$ for all $j$. This implies $\eta = \alpha \frac{\overline{\mu}}{q_s}$. As a result, the DM output level solves

$$\frac{\alpha^{1/(\eta - 1)} u'\left[\alpha^{1/(\eta - 1)} \mu q_s\right]}{c'(q_s)} = \frac{\eta}{\eta - 1} (1 + i)$$

(22)

I define equilibrium for the model with entry.

**Definition 2.** An equilibrium is a list $(p, q_s, \mu)$ that solves (10) (modified with $\sigma = 1$), (11), and (21).

I reduce equilibrium to a pair of equations for $(q_s, \mu)$:

$$\Gamma(q_s) = \frac{\eta}{\eta - 1}$$

(23)

$$\frac{\alpha^{1/(\eta - 1)} u'\left[\alpha^{1/(\eta - 1)} \mu q_s\right]}{c'(q_s)} = \frac{\eta}{\eta - 1} (1 + i)$$

(24)

Equation (23) says that in equilibrium average cost is a constant markup over marginal cost. This determines $q_s$. Equation (24) describes a markup relationship that determines $\mu$ given $q_s$. Equilibrium has a recursive structure. I first determine $q_s$ from (23) and then determine $\mu$ from $q_s$ and (24).
Figure 2 illustrates equilibrium with two graphs. The first plots average cost against marginal cost and indicates the equilibrium point $q_e$ together with the efficient scale $q_f$. The second graph takes $q_s$ as given and plots cost of holding real balances $\eta^{-1}(1+i)c'(q_s)$ against marginal utility as a function of the measure of sellers. I observe that the amount each firm produces $q_s$ is given uniquely by $\Gamma(q_s) = \frac{\eta}{\eta - 1}$. It is independent of search frictions. Let $q_f$ define the efficient scale, where average cost is minimized. In particular, for greater taste in variety, the gap $(q_f - q_s)/q_f$ is higher.

Comparing (23) and (20), we see that as $\epsilon[\alpha(\mu)] \leq 1$, the ratio of average to marginal costs at the social optimum is less than or equal to the value at equilibrium. This implies $q^*_s > q^*_e$; Equality only holds if $\alpha(\mu) = \mu$. Note that in the general linear case $\alpha(\mu) = A + B\mu$, with the constraint that $0 < A \leq \mu(1-B)$, we have $\epsilon(A + B\mu, \mu) < \epsilon(B\mu, \mu) = 1$.

**Lemma 3.** $1 < \Gamma(q^*_s) \leq \Gamma(q^*_e)$, so that $q^*_f > q^*_s \geq q^*_e$, with equality holding if and only if $\alpha(\mu) = \mu$.

The special case $\alpha(\mu) = \mu$, along with $i = 0, c(\cdot) = c$, corresponds to DS. DS show that firm output is the same between the unconstrained optimization problem of the social planner and equilibrium. Hence search frictions break the equivalence.

To analyze (24), it is helpful to formally define the marginal utility of a particular variety as a function of the measure of active sellers. For fixed $q_s$, let $\lambda(\mu) = \alpha(\mu)^{1/(\eta-1)}u''[\alpha(\mu)^{1/(\eta-1)}\mu q_s]$. $\lambda(\mu)$ captures the marginal utility of a particular variety as a function of the measure of sellers, where the production of each variety is fixed at $q_s$. Increasing the measure of sellers has two effects. The bundle of goods $\tilde{q}$ increases, so that the marginal utility thereof decreases. Second, $\alpha(\mu)^{1/(\eta-1)}$, which is the rate of change of the composite good with respect to the individual good, increases. Alternatively stated, the quantity per variety is fixed, so that as the measure of sellers increases, buyers consume a greater overall quantity of goods but also greater variety of goods. The first effect reduces the marginal utility of a good and the second increases it. Which effect dominates depends on the elasticity of marginal utility, the elasticity of the matching function, and the elasticity of substitution. The exact relationship is given by

$$\epsilon[\lambda(\mu)] = \frac{1}{\eta - 1}\epsilon[\alpha(\mu)] + \epsilon \left[ u''[\alpha(\mu)^{1/(\eta-1)}\mu q_s] \right] \left[ \frac{1}{\eta - 1}\epsilon[\alpha(\mu)] + 1 \right]$$

(25)

I examine $\lambda(\mu)$ in the case of CRRA preferences.
Example 1. Let \( u(\cdot) = q^{1-\varepsilon}/(1 - \varepsilon) \). Then \( \lambda(\mu) = \alpha(\mu) \frac{1 - \varepsilon}{\eta - 1} \mu^{-\varepsilon} q_s^{-\varepsilon} \). Hence, \( \epsilon[\lambda(\mu)] = \frac{1 - \varepsilon}{\eta - 1} \epsilon[\alpha(\mu)] - \varepsilon \). Since \( \epsilon[\alpha(\mu)] \) is bounded above by 1, \( \epsilon[\lambda(\mu)] \leq \frac{1 - \eta \varepsilon}{\eta - 1} \).

I state this and some limiting results in a lemma.

Lemma 4. Let \( u(q) = q^{1-\varepsilon}/(1 - \varepsilon) \) for \( 0 < \varepsilon < 1 \). If \( \eta \varepsilon > 1 \) then \( \lambda'(\mu) < 0 \). As \( \eta \to \infty \), \( \epsilon[\lambda(\mu)] \to -\varepsilon \), the elasticity of marginal utility of consumption. As \( \varepsilon \to 0 \) (utility becomes linear in the composite good), \( \epsilon[\lambda(\mu)] \to \epsilon[\alpha(\mu)]/(\eta - 1) \).

In words, as the measure of sellers increase, there is diminishing marginal utility for the marginal good unless there is sufficiently high taste for variety and/or a high enough elasticity of the marginal utility of the composite good.

If \( \eta \varepsilon < 1 \), then the behavior of the matching function depends on \( \mu \), as the following example shows.

Example 2. If \( \alpha(\mu) = \mu/(1 + \mu) \), then \( \epsilon[\lambda(\mu)] = \frac{1 - \varepsilon}{\eta - 1} \frac{1}{1 + \mu} - \varepsilon \), which is positive for \( \mu \) sufficiently low if \( \eta \varepsilon < 1 \).

I state results on comparative statics of equilibrium:

Proposition 5. The comparative statics are provided by Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( q_s )</th>
<th>( p )</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>0</td>
<td>0</td>
<td>↓</td>
</tr>
<tr>
<td>( k )</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>( \eta )</td>
<td>↑</td>
<td>−</td>
<td>↓</td>
</tr>
</tbody>
</table>

Proof. I examine the comparative statics of equilibrium. Consider an increase in \( k \). Equation (23) shows that the quantity sold by sellers to all consumers, \( q_s \), is increasing. Hence, \( p = \frac{k + c(q_s)}{q_s} \) is increasing as well. Consider now an increase in \( \eta \). This implies a lower markup. In order for the right hand side of (23) to remain constant, \( q_s \) must rise. As we have seen, a rise in \( q_s \) does not imply a fall in \( \mu \). To see the effect on \( p \), write \( p = \frac{k + c(q_s)}{q_s} \), for which

\[
\frac{\partial p}{\partial q_s} = \frac{q_s c'(q_s) - [k + c(q_s)]}{q_s^2}
\]

which is positive for \( k \) sufficiently low and negative for \( k \) sufficiently high. Hence, it is ambiguous.

Finally, suppose \( i \) increases. Then the left hand side of (24) must increase. \( q_s \) is determined independently of \( i \), so that \( \mu \) must adjust. As we have seen, the effect on \( \mu \) is ambiguous.

To complete the analysis with respect to \( \mu \) we consider the cases \( \lambda'(\mu) > 0 \) and \( \lambda'(\mu) < 0 \). Suppose \( \lambda'(\mu) < 0 \). Then \( \mu \) falls whenever the cost of holding real balances increases. As already shown, increases in \( i \) or \( k \) shifts the cost of holding real balances up. Suppose instead that \( \lambda'(\mu) > 0 \). Then \( \mu \) increases whenever the cost of holding real balances decreases.
By (23), (24), necessary conditions for equilibrium to be efficient are

\[ \epsilon[\alpha(\mu)] = 1 \quad (26) \]
\[ \frac{\eta}{\eta - 1} (1 + i) = 1 \quad (27) \]

Equation (26) says that the elasticity of the matching rate with respect to the sellers equals unity. This is a version of the Hosios condition (1990), which balances the thick market and congestion externalities. The entry of a new firm increases the variety of goods for consumers but reduces the buyers of other firms.

As previously discussed, the matching elasticity is strictly bounded above by unity, so the Hosios condition cannot be satisfied. Similarly, as \( \eta > 1 \), the second necessary condition is inconsistent.

**Proposition 6.** Every equilibrium is inefficient.

These inconsistent systems, however, make it clear that the only way a deviation from the Friedman rule could be welfare improving is by reducing the measure of sellers and thereby raising the elasticity of the matching function. However, we show in general that the Friedman rule maximizes equilibrium welfare.

**Proposition 7.** Let \((q_s, \mu)\) be an equilibrium. Then welfare is maximized at \(i = 0\).

**Proof.** Consider a social planner who takes \(q_s\) as given and chooses \(\mu\) to maximize social welfare. Then we check that the corresponding \(i\) from (24) is never satisfied with \(i > 0\). The planner solves

\[ \max_{\mu} \left\{ u\left(\frac{\alpha^{1/(\eta - 1)} \mu q_s}{\eta - 1}\right) - \mu [c(q_s) + k] \right\} \]

given \(q_s\) satisfying (23). The first order condition is

\[ u'[\alpha^{1/(\eta - 1)} \mu q_s]\left[\frac{1}{\eta - 1} \alpha^{(2 - \eta)/(\eta - 1)} \alpha'(\mu) \mu + \alpha(\mu)^{1/(\eta - 1)}\right] = \frac{c(q_s) + k}{q_s} \quad (28) \]

In general, (28) is not sufficient. Let \(\mu^*\) denote the welfare maximizing root. Using (24) and (23) we write

\[ \frac{c(q_s) + k}{q_s} = \frac{\alpha(\mu^*)^{1/(\eta - 1)} \mu' \left(\alpha(\mu^*)^{1/(\eta - 1)} \mu^* q_s\right)}{1 + i} \]

Hence we can rewrite (28) in terms of the interest rate as

\[ (1 + i) \left[\frac{1}{\eta - 1} \alpha(\mu^*)^{(2 - \eta)/(\eta - 1)} \alpha'(\mu^*) \mu^* + \alpha(\mu^*)^{1/(\eta - 1)}\right] = \alpha(\mu^*)^{1/(\eta - 1)} \]

This simplifies to

\[ 1 + i = \frac{\eta - 1}{\epsilon[\alpha(\mu^*)]} + \frac{\eta - 1}{\eta - 1} < 1 \]

This is a contradiction because \(i \geq 0\). This of course results from the fact that we considered a social planner which can control the measure of sellers directly, whereas the monetary authority can only affect the measure of sellers by varying the nominal interest rate. The relevant implication is that \(i = 0\) maximizes social welfare.

The optimality of the Friedman rule contrasts with competitive equilibrium in RW. The essential difference is that in RW there is a positive probability of sellers paying a fixed cost to enter and not matching with any buyers. This results in marginal cost exceeding average cost in equilibrium, so that firms operate beyond...
efficient scale even with no taste in variety. Reducing sellers reduces probability of non-trade and brings trade closer to efficient scale, so inflation may be useful. Hence, the details of search frictions can be important with respect to the welfare properties of inflation.

5.2.1. Non-monotonicity of consumer surplus

From the basic model, we know that the change in consumer surplus is given by $-\mu(\sigma + i)\mu_p c''(q_s)q_s^\varepsilon$. Since $\Gamma(q_s^\varepsilon) = \mu_p$, we can rewrite this as

$$\Omega'(q_s^\varepsilon) = -\mu c'(\sigma + i)\mu_p c''[\Gamma^{-1}(\mu_p)]\Gamma^{-1}(\mu_p)$$

(29)

Noting that $\Gamma^{-1}(\mu_p) = \frac{1}{\Gamma(q_s^\varepsilon)} > 0$, and $\Omega'(q_s^\varepsilon)$ is more negative, provided $\lambda'(\mu^\varepsilon) < 0$. Higher markups increase $q_s$, which interact with convex costs to make the change in consumer surplus more negative unless the measure of sellers decreases sufficiently. The behavior of consumer surplus is sensitive to firm entry in two ways: it depends directly on the measure of sellers, and it depends on quantities provided by each firm given by (23).

5.2.2. Equilibrium types: single crossing, no crossing, double crossing

Equilibrium is not in general unique, as it depends on the crossing of $\lambda(\mu)$ and the cost of holding real balances. I highlight the possibilities of uniqueness, nonexistence, and multiplicity using the same functional forms as Example ?? with varying parameter values:

Figure 3: Uniqueness, nonexistence, and multiplicity

With $\eta \varepsilon > 1$, $\lambda(\mu)$ slopes downward in the subplot (a). There is a single crossing between $\lambda(\mu)$ and the cost of holding real balances. Subplots (b) and (c) consider $\varepsilon = 0.25, \eta = 3$ so that $\lambda(\mu)$ is initially increasing. This means that the taste for diversity is stronger and that that marginal utility is less elastic. Here $\lambda(\mu)$ does not cross the adjusted price curve for the given costs of entry. In (c) preferences are the same, but costs of entry are lower, so that there is a double crossing of $\lambda(\mu)$ and adjusted prices. The two equilibria for these particular parameter values are given by $(q_s, \mu) = (0.556, 0.133)$ and $(0.556, 1.591)$. In Appendix A.3, I show that the second derivative of the buyer’s objective function is negative at these two values, so the first order condition is sufficient.

I thus have a low-variety and a high-variety equilibrium in Figure 3 (c). In the low variety equilibrium, there are not too many sellers, which keeps demand low given the matching technology. Given the low demand, sellers do not have an incentive to enter. In the high variety equilibrium, however, the taste for

---

13Specifically, in RW, the free entry condition is $a_s(n)[c'(q_s)q_s - c(q_s)] = k$, which can be rearranged as $\frac{k}{q_s c'(q_s)} = 1$, so that average costs exceed marginal costs.
variety induces a higher demand, which is enough to sustain a greater measure of sellers. Note that we can compare social welfare between the low-variety equilibrium and high-variety equilibrium by comparing
\[ u[\alpha(\mu)^{\frac{1}{\eta-1}} \mu q_s] - \mu \mu_{p} c'(q_s). \]
Social welfare turns out to be higher in the second case. Since sellers are indifferent in either equilibrium due to the zero profit condition, the low-variety equilibrium is Pareto inferior to the high-variety equilibrium. Thus, there is a coordination problem. If a sufficiently great mass of sellers entered, the extra variety would push up marginal utility and raise demand enough for the sellers to break even. However, an individual seller has no incentive to enter unilaterally. However, this multiplicity requires very low elasticity of substitution.

**Theorem 1** (Existence and Uniqueness). Equilibrium exists if \( \lim_{\mu \to 0} \lambda(\mu) \to \infty \) and \( \lim_{\mu \to \infty} \lambda(\mu) \to 0 \). Equilibrium is unique if \( \lambda'(\mu) < 0 \) for all \( \mu \).

**Proof.** We have already shown that there is a unique \( q^*_s \) satisfying \( \Gamma(q^*_s) = \frac{\eta}{\eta-1} \). Given \( q^*_s \), let \( P = (1 + i) \frac{\eta}{\eta-1} c'(q^*_s) \), the price adjusted for the cost of holding real balances. It suffices from (24) to show that there exists \( \mu \) such that \( \lambda(\mu) = P \). The two Inada conditions on \( \lambda(\mu) \) ensure that there exists \( \mu^*, \mu^{**} > 0 \) for which \( \lambda(\mu) > P \) for \( \mu < \mu^* \) and \( \lambda(\mu) < P \) for \( \mu > \mu^{**} \). By continuity, there exists \( \mu^{***} \) such that \( \lambda(\mu^{****}) = P \). Moreover, if \( \lambda(\mu) \) is decreasing everywhere, then equilibrium is unique.

It is easy to see that both conditions hold with CRRA preferences and \( \eta \varepsilon \geq 1 \).

**Corollary 2.** Given \( u(q) = q^{1-\varepsilon}/(1 - \varepsilon) \), there is a unique equilibrium if \( \eta \varepsilon \geq 1 \).

5.3. **Scale economies and comparison with Dixit-Stiglitz**

DS focuses on scale economies with the simplifying assumption of constant marginal costs and decreasing average costs. Their model is a special case of the model with entry here (\( \alpha(\mu) = \mu \)). DS finds that a constrained optimum coincides with the equilibrium in output and number of firms, and that an unconstrained optimum has the same output per firm but more output overall and hence more firms. The social planning problem we considered here is an unconstrained one. I already showed \( q^*_s \geq q^*_e \), with equality holding when \( \alpha(\mu) = \mu \). I now further assume \( c(\cdot) = c \) and explicitly compare the equilibrium and unconstrained optimization.

\[
\left(\frac{\alpha(\mu^*)}{\alpha(\mu^e)}\right)^{1/(\eta-1)} \frac{u'(\bar{q}^*)}{u'(\bar{q}^e)} = \frac{1}{\eta-1} \frac{u'(\bar{q}^*)}{u'(\bar{q}^e)} < 1
\]

Equation 30 and Lemma3 imply

**Proposition 8.** For sufficiently high \( \eta \), elasticity of the matching function, or \( i, \mu^* > \mu^e \).

Thus, the DS result of a greater measure of sellers in the social optimum generalizes to a monetary economy when search frictions and markups are low but does not hold generally. Otherwise, there are too many firms, or too much variety.

We can characterize this more closely with CRRA preferences.

---

14 The caveat is that they have general homothetic preferences over the monopolistic sector and outside good rather than quasilinear utility.

15 The constraint considered by DS is that no lump-sum subsidies can be provided to monopolistic firms.
\[
\frac{\left[\alpha(\mu^*)\right]^{\frac{1-\epsilon}{\eta-1}}}{\left[\alpha(\mu^e)\right]^{\frac{1-\epsilon}{\eta-1}}} = \left(\frac{q_s^e}{q_s^*}\right)^{\epsilon} \frac{1}{\mu_p(1+i)}
\]

(31)

The comparison between \(\mu^*\) and \(\mu^e\) depends on whether the right hand side exceeds 1. Suppose \(\eta\varepsilon > 1\). If the right hand side exceeds 1 then \(\mu^* > \mu^e\). Otherwise, \(\mu^* < \mu^e\). In the frictionless benchmark, \(\alpha(\mu) = \mu\), and using \(q_s^e = q_s^*\) we obtain

\[
\frac{\mu^*}{\mu^e} = [\mu_p(1+i)]^{\frac{\eta-1}{\eta-1}}
\]

In the case of frictionless matching, \(\eta\varepsilon > 1\) implies \(\mu^* > \mu^e\). With search frictions, it depends on whether the right hand side exceeds 1. There is no general way to bound \(q_s^*/q_s^e\) because it becomes arbitrarily large as \(\frac{\epsilon[\alpha(\mu)]}{\eta-1} \to 0\), which can happen if the elasticity of the matching function approaches zero or as goods become perfectly substitutable. In that case, it is socially optimal to have very few sellers producing many goods.

Figure 4: Comparison of measure of sellers

\[
\Gamma(q_s^*) = \frac{\epsilon[\alpha(\mu)]}{\eta-1} + 1
\]

\[
\Gamma(q_s^e) = \frac{\eta}{\eta-1}
\]

Figure 5 shows the two possibilities for the case in which \(\lambda(\mu)\) is diminishing with respect to \(\mu\). Abusing notation, let \(\lambda(\mu, q_s^e), \lambda(\mu, q_s^*)\) denote \(\lambda(\mu)\) for \(q_s\) equal to the equilibrium and socially optimal amounts, respectively. The first plot has relatively high nominal interest rate and relatively few search frictions, so that \(\mu^* > \mu^e\). The opposite is true in the second plot, so that \(\mu^e > \mu^*\). Figure 6 depicts the frictionless case: \(\alpha(\mu) = \mu\).
In this case, there is no congestion externality, but there is reduced demand from both the inflation wedge and price markups. Hence, $\mu^* > \mu^e$. The case considered in DS arises simply from setting $i = 0$, in which case price markups would be the only inefficiency.

6. Variable markups and the role of complementarity

I emphasized CES preference because of simplicity and to facilitate comparison with Dixit Stiglitz. But there are well known problems with CES. Zhelobodko, Kokovin, Parenti, and Thisse (2012), hereafter ZKPT, identify two important deficiencies: (1) markups and prices are independent of firm entry and market size and (2) the lack of a scale effect (the size of firms and markups are independent of the number of consumers). The free entry of sellers makes (1) particularly salient in this context. Furthermore, the marginal utility of a variety $u'(q)\alpha(\mu) \overline{\nu - 1}$ decreases with fewer sellers because of complementarity, yet there is no change in markup. We investigate the importance of these channels by relaxing complementarity with additively separable preferences and introducing variable markups. This enables us to decompose the change in welfare costs of inflation into (1) complementarity effects and (2) variable markup effects.

Variable markups require changing the preferences, and we adopt additively separable preferences over varieties a la ZKPT in the DM.

6.1. The buyer’s problem

The preferences of the buyer in the DM are given by

$$\int_0^\Omega u(q_i)di = \int_0^\Omega u(q_i)di + U(x) - h$$

where $u(0) = 0$, $u'(0) = \infty$, $u'(q) > 0$, and $u''(q) < 0$ for $q > 0$. $U(\cdot)$ satisfies the same conditions as $u(\cdot)$. 
The concavity of \( u(\cdot) \) reflects taste for variety. Consumers prefer to spread consumption over all varieties than a small mass of varieties. There is a formal equivalence between decision-making by consumers with taste for variety and the Arrow-Pratt theory of risk aversion. Consumers’ taste for variety can be measured from the relative love for variety (RLV)

\[
    r_u(q) = -\frac{q u''(q)}{u'(q)} > 0
\]  

which is the familiar elasticity of marginal utility, or inverse of elasticity of substitution in the case \( q_i = q \quad \forall i \). Preferences which display an increasing RLV mean that consumers perceive varieties as being less substitutable when they consume more. Preferences may also display a decreasing RLV and be more substitutable with higher consumption. We assume \( r_u'(q) < 2 \quad \forall q > 0 \), which we shall see makes the producers’ problem concave.

Figure 6 depicts the RLV for several functional forms.

The problem for the buyer is given by

\[
    \max_{q_i} \left\{ \int_0^{\alpha(\mu)} u(q_i) di - (1 + i)z \right\}
\]

where \( z = \int_0^{\alpha(\mu)} p_i q_i di \) as before. The solution is given by

\[
    u'(q_i) = (1 + i)p_i
\]

where an interior solution is guaranteed because of the Inada condition. \( p_i(q_i) \) is strictly decreasing because \( u(\cdot) \) is strictly concave. The elasticity of the inverse demand \( \varepsilon_p(q_i) \) and the price elasticity of demand are related to the RLV as

\[
    \frac{1}{\varepsilon_p(q)} = \varepsilon_p(q) = r_u(q)
\]

Hence, the price elasticity of demand is just the reciprocal of the RLV. This implies that RLV increases if and only if demand for a variety becomes less elastic with quantity (more elastic with price). Intuitively, consumers are less willing to substitute goods with higher consumption. Hence, as with CES, taste for variety, elasticity of substitution, and price elasticity of demand are interchangeable, but in contrast to CES they depend on the consumption level of \( q_i \).\textsuperscript{16}

\textsuperscript{16}The preceding is a special case of the discussion in Mrazova and Neary (2013). Elasticity of demand decreases with sales
Furthermore, suppose that there are more available varieties, and consumers spread out consumption among the greater set of varieties. Then \( q \) is lower for each variety, so that with increasing RLV, \( r_u(q) \) is lower. Hence, there is more substitutability between varieties. This confirms the intuition that the substitutability of varieties increases with the number of varieties, all else constant.

6.2. The producers’ problem

The producer \( j \) produces \( q_j \) for measure \( \alpha_s = \frac{\alpha(q)}{\kappa} \) consumers. As before, let \( q_s(j) = \alpha_s q(j) \)
\[
\max_{p_j, q_j} \{ p(j) q_s(j) - c(q_s) \}
\]
where \( p(j) \) is given by (36). The solution equates marginal revenue \( \frac{u'(q) [1 - r_u(q)]}{1 + i} \) to marginal cost \( c'(q_s) \), and yields
\[
r_u(q_i) = \frac{u'(q_i) - (1 + i) c'(q_s)}{u'(q_1)}
\]
where \( r_c = -qc''/c' \) is the negative elasticity of marginal cost. By construction, \( r_c < 0 \), so that the second term in (40) is positive. \( r_u(q_i) \) measures the convexity of demand for \( q_i \), so the second order condition requires that demand not be too convex. \( r_u'(q_i) < 2 \) \( \forall q \geq 0 \) is sufficient because the second term is nonnegative. Note that, in the CES case \( u(q) = q^{\frac{1}{\kappa}} \), \( r_u = \varepsilon \), and \( r_u' = 1 + \varepsilon \), so that \( \varepsilon < 1 \) is a sufficient condition.

The condition that marginal revenue equals marginal cost can be expressed as \( p(q_s) + q_s p'(q_s) = c'(q_s) \), which implies that \( \frac{p - c'(q_s)}{p(q_s)} = -\frac{q_s p'(q_s)}{p(q_s)} = \frac{1}{r_u(q_s)} \). Hence, an increase in marginal cost \( c \), which must lower sales, is associated with a higher elasticity. It is useful to define the curvature of demand \( \xi(q) = -\frac{q p''(q)}{p'(q)} \). Following Mrazova and Neary (2013), by differentiating the firm FOC with respect to \( c'(q_s) \), \( \frac{dp}{dc'} = \frac{1}{2\xi - 1} \), which implies \( \frac{dp}{dc'} - 1 = \frac{\xi - 1}{2\xi} \). Hence, there is more than 100% pass-through of costs to prices if and only if \( \xi > 1 \).

As the model with entry, producers enter in the CM at cost \( k \) up to the point until profits are zero:
\[
k + c(q_s) = p(j) q_s
\]
which satisfy
\[
\Gamma(q_s) = \frac{1}{1 - r_u(q)}
\]
\[
u'(q) = \frac{1 + i}{1 - r_u(q)}
\]
\[
c'(q_s) = \frac{c'(q_s)}{r_u(q)}
\]
\[
p = \frac{1}{1 - r_u(q)}
\]
\[
q_s = \frac{\alpha(q)}{\kappa} q
\]
if and only if the inverse demand \( p(q) \) is subconvex, which means that \( \log p(q) \) is concave in \( \log q \). Alternatively, demand is superconvex if elasticity of demand increases with sales. These notions will be important for the producers’ problem.
Since \( r_u(q) \) is also the net markup, this means that the ratio of average to marginal costs equals the gross markup in equilibrium.

Equilibrium can be described in \((p, q)\) space in terms of the following equations:

\[
p = \frac{u'(q)}{1 + i} 
\]

(46)

\[
p = \frac{c'[\phi(q)]}{1 - r_u(q)} 
\]

(47)

(48)

where \( \phi(q) = \Gamma^{-1} \left[ \frac{1}{1 - r_u(q)} \right] \) and satisfies \( \phi'(q) < 0 \). Equation (46) is a demand curve, which depends on the cost of holding money, and Equation (47) is price setting rule, which reflects the markup directly and also indirectly via the effect on marginal costs expressed by \( c'[\phi(q)] \).

Existence and uniqueness of equilibrium holds generally.

**Proposition 9.** Equilibrium exists and is unique.

**Proof.** First, we show that there is a unique crossing of demand and the price setting rule. It suffices to show that \( \frac{u'(q)(1 - r_u(q))}{(1 + i)c'[\phi(q)]} = 1 \) for unique \( q \). The denominator is increasing and strictly bounded below by zero. The numerator approaches zero because \( \lim_{q \to \infty} u'(q) = \infty \) and \( r_u(q) < 1 \) from firm optimization. This defines unique values \( p^* \) and \( q^* \). In turn, \( q_s^* = \phi(q^*) \) and \( \mu^* \) is defined implicitly from \( q_s^* = \frac{\alpha(\mu^*)}{\mu^*} q^* \).

Note that, in contrast to CES preferences, the Inada conditions on \( u(\cdot) \) suffice for existence and uniqueness because there is a lack of strategic complementarity of varieties: having more varieties does not increase the marginal utility of a particular variety. Hence, the marginal utility of a variety approaches zero as the number of sellers approaches infinite without any further assumptions.

We show that higher interest rates lead to lower overall consumption and larger and fewer firms.

**Lemma 5.** If \( r_u'(q) > 0 \) in the neighborhood of equilibrium, then a small increase in \( i \) leads to higher \( q_s \), lower \( q \), lower \( \mu \), and lower markups.

**Proof.** An increase in \( i \) shifts demand downward and does not change the price setting rule, resulting in lower \( p \) and \( q \). Hence, markups \( r_u(q) \) are lower. \( q_s = \phi(q) \) is higher. The measure of sellers satisfies \( \frac{\alpha(\mu)}{\mu} = \frac{\alpha(q)}{q} \). The right hand side is higher, so that \( \mu \) is lower.
It may seem counterintuitive there can be lower markups with fewer firms. The reasoning is as follows. A higher nominal interest rate makes it more costly to use money, lowering consumption. Lower demand leads to lower sales and lower markups on those sales since the taste for variety is lower. Hence, fewer firms enter.

**Lemma 6.** If \( r_u(q) > 0 \) in the neighborhood of equilibrium, then a small decrease in \( k \) leads to higher \( q \), lower \( q_s \), higher \( \mu \), and higher markups.

*Proof.* Lower \( k \) implies a rightward shift of the supply curve. This can be seen from the fact that \( \Gamma \) is shifted down and hence \( \phi \) is shifted up, so that \( q_s = \phi(q) \) is lower (\( \phi \) is negative). Hence firms have lower marginal costs \( c'(\phi(q)) \). Demand is unaffected, so equilibrium occurs at higher \( q \) and lower \( p \). Thus, \( r_u(q) \) and markups are lower. \( \frac{q}{q} \) is lower, so more sellers enter.

\[
\text{Table 3: Comparative statics: } r_u(q) > 0
\]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( q_s )</th>
<th>( q )</th>
<th>( r_u(q) )</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \uparrow i )</td>
<td>( \uparrow )</td>
<td>( \downarrow )</td>
<td>( \downarrow )</td>
<td>( \downarrow )</td>
</tr>
<tr>
<td>( \uparrow k )</td>
<td>( \uparrow )</td>
<td>( \downarrow )</td>
<td>( \downarrow )</td>
<td>( \downarrow )</td>
</tr>
</tbody>
</table>

In Figure 8, we consider a rise from \( i = 0 \) to \( i = 0.13 \) for \( u(q) = \frac{q^{1-\epsilon}}{1-\epsilon} + Bq \), so that \( r_u(q) = \frac{\epsilon}{1+Bq^\epsilon} \). This leads to a sharp decrease in \( q \) and increase in \( q_s \), reducing welfare significantly.

**Figure 8:** A rise in the nominal interest rate

6.2.1. **Social welfare**

The social welfare function can be written as

\[
W(q_s, \mu) = \alpha(\mu)u(q) - \mu[k + c(q_s)]
\]  

(49)
which has first order conditions\footnote{For (51) we obtain \( \alpha'(\mu)u(q) + u'(q)[1 - \alpha(\mu)]q_s = k + c(q_s) \) using \( \frac{\partial q_s}{\partial \mu} = \frac{1 - \alpha(\mu)}{\alpha'(\mu)} \). Then we divide both sides by \( q_s c'(q_s) \) and use the fact that \( u'(q) = c'(q_s) \) and rearrange in terms of \( \varepsilon_\alpha, \varepsilon_u, \) and \( \Gamma \).}

\[ u'(q) = c'(q_s) \quad (50) \]

\[ \Gamma(q_s) = \varepsilon_\alpha(\mu) \left( \frac{1}{\varepsilon_u(q)} - 1 \right) + 1 \quad (51) \]

This is a generalization of a result in Vives (1999). Without search frictions, note that \( \Gamma(q_s) = \varepsilon_u(q) \), which is the result in Vives (1999). More generally, the optimal deviation of average costs from the efficient scale increases with a more concave utility function and a less concave matching function.

Comparing (50)-(51) to (42) and (43), we find the one-way comparison \( q_s^* < q_e^* \Rightarrow q_s^* > q_e^* \) and hence \( \frac{q_s^*}{q_e^*} < \frac{q_s}{q_e} \Leftrightarrow \mu^* > \mu^e \). Thus, as in DS with variable elasticity of demand, if optimal firm size is smaller then the optimal number of firms is greater. Note that the relationship between \( q_s^* \) and \( q_e^* \) can go either way, because the former depends on the elasticity of utility and the elasticity of the matching function whereas the latter depends on the elasticity of demand. Compared to DS, however, \( q_e^* \) is higher if \( i > 0 \) because of the inflation wedge and \( q_s^* \) is lower because the social planner takes into account search frictions.

**Proposition 10. Equilibrium is inefficient.**

**Proof.** From (50)-(51) and (46)-(47), necessary and sufficient conditions for equilibrium to implement the social optimum are \( \frac{1+i}{1-r_u(q)} = 1 \) and \( 1 - r_u(q) = \varepsilon[u(q)]/\varepsilon[\alpha(\mu)] \). This requires \( i = 0, r_u(q^*) = 0 \) and \( \varepsilon[u(q^*)] = \varepsilon[\alpha(\mu^*)] \).

But \( r_u(q) = 0 \) is inconsistent with firm optimization. Hence, equilibrium is inefficient. \( \blacksquare \)

However, the Friedman rule is not necessarily optimal, as Figure 9 demonstrates.

\[ \text{Figure 9: Social welfare} \]

In this example, social welfare is maximized at \( i = 0.0173, 0.0361, \) and 0.0303 across subpanels a),b), and c), respectively. There are two benefits of inflation in equilibrium. It reduces price markups, and it reduces average costs of production. The reduction of price markups mitigates the inefficiency on the intensive margin. Higher inflation reduces product variety, but this is less important for welfare when product variety is already very high. A lower level or higher elasticity of costs shift the optimal nominal interest rate higher.

It is instructive to compare this optimal deviation from the Friedman rule to that of competitive equilibrium in RW. Inflation there reduces sellers and hence reduces congestion, which in turn lowers average costs by increasing the probability that sellers match with buyers. Here, lower congestion does not reduce average costs directly. Instead, average costs decrease because markups decrease. The reduction in average costs here rests on variable elasticity of demand, whereas in RW it results from the matching technology.
Inflation hence reduces sellers’ market whenever \( r_u \) is increasing. The idea that inflation can reduce sellers’ market power is not new. In Diamond (1993) or in a basic new Keynesian model, inflation reduces markups with sticky prices, but here the result instead results from changes in equilibrium taste of variety and without any nominal rigidity.

The ratio of average costs to marginal costs in the social optimum and equilibrium are given by \( \Gamma(q^*_s) = \frac{\varepsilon[\alpha(u)]}{\varepsilon[u(q) - r_u]} \) and \( \Gamma(q_e) = \frac{1}{1-r_u(q)} \). In words, the socially optimal quantity is uniquely determined by the elasticity of utility and elasticity of the matching, whereas the equilibrium quantity is determined by the elasticity of demand (or equivalently the RLV). In general, firms can be either too big or too small relative to the social optimum.

6.2.2. The elasticity of the markup and augmented HARA preferences

Taking \( r_u(q) = -\frac{u''(q)q}{u'(q)} \), applying logs and differentiating, we obtain the elasticity of the markup:

\[
\varepsilon_{r_u} = 1 + r_u(q) - \xi(q)
\]

where \( \xi(q) \leq 2 \). This says that the elasticity of the markup equals the gross markup minus the curvature of demand. Since \( r_u(q) = \frac{1}{\varepsilon[u(q)]} \), the elasticity of the markup is the negative elasticity of the elasticity of demand. One major task for calibration and estimation is the identification of a suitable choice of utility functions. In the working paper by ZKPT (2012), the authors consider an extension of HARA preferences that is consistent with both increasing and decreasing taste of variety. They dub it ‘augmented HARA’:

\[
u(q) = \frac{1}{\rho}[(a + hq)^\rho - a^\rho] + bq, \text{ where } a \geq 0, h \geq 0, b \geq 0, 0 < \rho < 1. \]

CES arises with \( a = b = 0 \), and HARA arises with \( b = 0 \). Furthermore, \( a > 0 \) bounds the marginal utility at zero consumption.

RLV takes the form \( r_u(q) = \frac{h^2(1-\rho)(a + hq)^{\rho-2}q}{h(a + hq)^{\rho-1} + b} \). With \( a = 0 \), as we wish to include an Inada condition, this becomes \( r_u(q) = \frac{(1-\rho)hq^{-1}}{hq^{-1} + b} \). A crucial question regards the behavior of the elasticity of the taste of variety (elasticity of markups in equilibrium). Denoting this by \( \varepsilon_{r_u} \), we have

\[
\varepsilon_{r_u} = \frac{(\rho - 1)b}{\rho hq^{-1} + b}
\]

For augmented HARA preferences, \( \xi(q) = \frac{(2-\rho)hq}{a + hq} \), which is positive and increasing unless \( a = 0 \). At \( a = 0 \), this simplifies to \( 2 - \rho \). Hence, at \( a = 0 \), \( \varepsilon_{r_u} = \rho - 1 + r_u(q) \), which is increasing with \( q \) for \( r_u > 0 \).

7. Optimality of Friedman rule

Table 4 summarizes the optimality properties of the Friedman rule under various market structures. Here, ‘first best’ refers to the implementation of the social planning problem, and ‘second best’ refers to the maximization of equilibrium welfare. For the first four market structures, I refer to the analysis in Rocheteau Wright (2005). The next two are from the variety model in Dong (2010). The last three are the subject of this paper.

\[^{18}\text{If } b = 0, \text{ the coefficient of absolute risk aversion satisfies } \frac{u''(q)}{u'(q)} = \frac{h(\rho - 1)}{1 + hq}, \text{ which is hyperbolic. Furthermore, } r_u(q) \text{ is decreasing for } a > 0 \text{ and constant for } a = 0. \text{ This explains the need to modify HARA to be consistent with increasing RLV.}\]
Table 4: Optimality of Friedman rule

<table>
<thead>
<tr>
<th>Market structure</th>
<th>First best</th>
<th>Best policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Nash bargaining</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Proportional bargaining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Price taking</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Price posting</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Dong (Nash)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Dong (price posting)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MC fixed sellers</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>MC entry of sellers CES</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>MC entry of sellers ZKPT</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

8. Measuring the welfare costs of inflation

8.1. The compensated measure of the welfare cost of inflation

I estimate the money demand implied for each model with the empirical money demand from the United States. Figure 10 compares the predicted money demand curve among the basic model, the model with free entry, and the baseline LW.

Figure 10: Money demand

Following Lucas (2000), we set period length to a year and \( \beta^{-1} = 1.03 \). Given the assumption that \( r = 3\% \) is consistent with zero inflation we ask: what is the percentage \( \Delta \) of total consumption that individuals would be willing to sacrifice in order to be in the steady state with an interest rate of 3\% instead of the steady
I use the money demand data to match the first four moments of money demand for different values of \( \eta \). Estimation is with respect to \((\varepsilon, A)\) in the basic model and \((\varepsilon, A, k)\) in the free entry model. We calculate the cost \( \Delta \) of 10% inflation, which corresponds to \( r = 0.03 \). The formulas for \( \Delta \) are given by the following. In the basic model, sellers make equilibrium profits, which we distribute back to the buyers without loss of generality because of quasilinear utility. With entry, equilibrium profits are zero. Furthermore, we rewrite the \( \alpha q \) units purchased by the buyer in the DM as \( \alpha^{-\frac{1}{1-\varepsilon}}q_{0.03}^{-1-\varepsilon} + A \ln(1-\Delta) - \frac{\eta}{\eta-1} \alpha^{-\frac{1}{1-\varepsilon}}q_{0.03}^{\varepsilon} \). For additively separable preferences, the welfare cost \( \Delta \) is given by

\[
\alpha(\mu_{0.03})u[q_{0.03}(1-\Delta)] + A \ln(1-\Delta) - \alpha(\mu_{0.03})q_{0.03} c'(q_{s,0.03}) = \alpha(\mu_{r})u(q_{r}) - \alpha(\mu_{r})q_{r} c'(q_{s,r})
\]
rationale is that inflation reduces entry, which in equilibrium reduces both variety and also reduces the value of existing varieties due to the complementarity effect.

At $\mu_P = 1.01$, taste of variety and markups are very low, so the economy resembles Walrasian price taking. It is instructive to compare this to Rocheteau Wright (2009), who found welfare costs of inflation at 1.54% without entry and 1.65% with entry. These values are slightly lower but comparable to the present case. Two important differences are that without markups they need to introduce some convexity in costs; specifically, they set $\delta = 1.1$ in $c(q) = q^\delta/\delta$, and that there is still some small taste for variety and market power in our case.

9. Conclusion

In this paper, I examined the implications of product variety and multilateral matching for firm size, variety, and the welfare costs of inflation. Markups amplify the costs of holding real balances through a rent sharing externality but also help induce entry. Firms produce too little under CES (with or without entry). Furthermore, with entry, there can either be too many or too few firms relative to the social optimum, and are likely to be too few with a higher nominal interest rate, a more elastic matching function, and lower taste for variety. With and without entry, the Friedman rule is best policy with CES preferences and attains first best as $\eta \to \infty$. The estimated welfare costs of inflation under CES are 7% without entry and over 9% with entry at 30% markups. This is much higher than previous estimates by Lucas (2000) and LW. It is in line with the highest estimates of RW (2009) and Dong under bargaining. Furthermore, the Friedman rule is not generally optimal with variable elasticity of demand because inflation reduces markups and, in turn, average costs. This effect also attenuates the inefficiency on the intensive margin. However, the Friedman rule is still optimal at parameter values obtained from matching money demand. In the appendix, I also extend the model to variable search intensity and show that the hot potato effect (rising search intensity to inflation) identified by LLW requires implausible parameter values if there is congestion in search. This, of course, does not negate the existence of the hot potato effect, but highlights the importance on the role of congestion on search intensity.

There are many fruitful directions for further research. For one, the estimates from the welfare cost of inflation derived from the model are likely upward biased from at least two different reasons. Notably, the framework rests on a pure monetary economy, but credit provides another means of payment and thereby mitigates the inflation wedge. Second, markups reasonably result from reasons besides taste for variety, such as imperfect intermediate goods markets (as in a canonical new Keynesian model). Hence, a reasonable extension is to modify the model with money and credit, with limited commitment or enforcement constraints, and have an intermediate goods sector. Furthermore, more general specifications of variable markups can be used to incorporate degrees of complementarity.
Appendix A. Appendix

Appendix A.1. Endogenous search intensity

Appendix A.1.1. Environment

I extend the model by endogenizing buyers’ search for sellers. The main motivation is to understand how shopping effort and output varies in response to interest rates and compare to LLW. LLW, also using a flexible labor market, find that with enough substitutability of labor and goods in preferences, inflation increases and labor, search effort, and output increase if taste for variety is high enough. In contrast to LLW, we consider the congestion effect of search. I find that though it is possible for effort to rise in response to interest rates, it requires the taste for variety to be implausibly high relative to the elasticity of the marginal utility of consumption. Typically, we will see that effort falls modestly.

The measure of overall matches is given by \(\alpha(\bar{\pi})\), where

\[
\bar{\pi} = \int_0^1 e(i)di \tag{A.1}
\]

The measure of varieties of the DM good that a buyer can purchase is given by \(\alpha_b = \alpha(\bar{\pi})/\bar{\pi}\). Since buyers can match up with at most measure \(\mu\) sellers, \(\alpha(\bar{\pi}) \leq \mu\). The measure of buyers for each seller is \(\alpha_s = \alpha(\bar{\pi})/\bar{\pi}\).

I assume \(\lim_{e \to \infty} \alpha_s(e) = 1\)

The aggregate good is a function of the measure of contacted sellers, and hence a function of effort

\[
\eta(\bar{\pi}) = \left(\int_0^{e\alpha/\bar{\pi}} q(j)\frac{u-1}{\eta} dj\right)^{\frac{\eta}{\eta-1}}
\]

Effort yields disutility given by \(\psi(e)\), where \(\psi'(e) > 0, \psi''(e) > 0\) and \(\psi(0) = \psi'(0) = 0\). The buyer’s period utility function becomes \(U^b(x, h, \eta, e) = u(q) - \psi(e) + U(x) - h\).

Though we do not explicitly introduce a labor-leisure tradeoff as LLW do, we proxy one via the convexity of \(\psi(e)\).

Appendix A.1.2. Equilibrium

The feasibility constraint on the transfer of real balances \(\int_0^{e\alpha/\bar{\pi}} p(j)q(j) dj \leq z\) is binding and the buyer’s problem is expressible as

\[
\max_{q(j), e} \left\{ -(1 + i) \int_0^{e\alpha/\bar{\pi}} p(j)q(j) dj - \psi(e) + u(\eta) \right\}
\]

The first order condition with respect to \(q(j)\) is

\[
(1 + i)p(j) = u'(\eta) \left[ \frac{\eta}{q(j)} \right]^{1/\eta} \tag{A.2}
\]

as before.

The first order condition with respect to effort is:

\[
u'(\eta) \left(\frac{\eta}{\eta-1}\right)^{1/\eta} q^{\frac{\alpha}{\bar{\pi}}} - \psi'(e) - (1 + i)p(j)q(j)\frac{\alpha}{\bar{\pi}} = 0
\]

This can be simplified using (A.2) into

\[
\psi'(e) = \frac{(1 + i)p(j)q(j)\alpha/\bar{\pi}}{\eta - 1} \tag{A.3}
\]
Equation (A.3) says that the marginal cost of effort is equal to the marginal utility of effort. Note that (A.3) reflects a ‘hot potato’ effect. A higher interest rate raises the cost of holding unused real balances and thereby raises search intensity among buyers. However, this is a partial equilibrium effect. I later show that in general equilibrium effort is likely to fall due to congestion.

The interpretation of (A.3) is as follows. An extra unit of effort enables the buyer to access \( \alpha/\bar{\tau} \) extra sellers and hence \( q \) extra units. These extra units cost \( (1+i)q(j) \) each, which equals their marginal benefit by (A.2). Since the right hand side is positive and does not depend on \( e \), and since \( \psi'(e) > 0 \) with \( \psi'(0) = \psi'(\infty) = \infty \), there is a unique value of \( e \) that solves (A.3). As all buyers face the same problem, \( e = \bar{\tau} \).

The problem of seller \( j \) is identical to the basic case except that now \( q_s(j) = \alpha_s(e)q(j) \). Similarly, as buyers face the same problem with a unique solution, \( q(j) = q \) for all \( j \) and \( \bar{q} = \alpha^{n/(\eta-1)}q \). I am now ready to define equilibrium.

**Definition 3.** An equilibrium for the model with endogenous search intensity is a list \((q_s, e)\) such that

\[
\frac{\alpha(e)^{1/(\eta-1)}u'[\alpha(e)^{1/(\eta-1)}\mu q_s]}{e'(q_s)} = \frac{\eta}{\eta-1}(1+i) \quad (A.4)
\]

\[
e\psi'(e) = \frac{\eta}{(\eta-1)^2}e'(q_s)\mu q_s(1+i) \quad (A.5)
\]

I prove existence of equilibrium.

**Theorem 2** (Existence and uniqueness). There exists a unique equilibrium.

**Proof.** Fix \( e \) in \((0, \infty)\). This fixes \( \alpha(e) \in (0, 1) \). By the standard argument, tending \( q_s \to 0 \) tends the left hand side of (A.4) to \( \infty \), and tending \( q_s \to \infty \) tends the left hand side to 0. By the intermediate value theorem, there is a (positive) \( q_s \) such that (A.4) holds. Given that the left hand side is decreasing in \( q_s \), this value is unique. By the implicit function theorem, this defines a differentiable function \( q_s = f(e) \) such that \( f'(e) < 0 \). Moreover, using the fact that as \( e \to \infty, \alpha/\mu \to 1 \), we have \( \lim_{e \to \infty} f(e) \) is given by the unique value \( q_s^\infty \). Also, \( \lim_{e \to 0} f(e) = \infty \).

From (A.5), define the function

\[
g(e) = e\psi'(e) - \frac{\eta(1+i)}{(\eta-1)^2}e'[f(e)]\mu f(e)
\]

It suffices to show that there is a unique \( e > 0 \) such that \( g(e) = 0 \). As \( f'(e) < 0, g'(e) < 0 \). Furthermore, as \( e \to \infty, f(e) \to q_s^\infty \), so that \( g(e) \to \infty \). Second, note that \( g(e) \to -\infty \) as \( e \to 0 \). As \( e \to 0, f(e) \to \infty \). By the intermediate value theorem, there is a unique \( e^* > 0 \) such that \( g(e^*) = 0 \). \( \square \)

I can describe equilibrium in terms of two curves relating \( e \) and \( q_s \). Equation (A.4) is the marginal markup condition. (A.5) describes the effort as a function of firm output. Specifically, there is a linear relationship between the semi-elasticity of effort cost and the semi-elasticity of production cost.

I derive a necessary and sufficient condition for effort to fall in equilibrium with respect to interest rates. I take the logarithmic transforms of (A.4) and (A.5), totally differentiate with respect to \( 1+i \), and rearrange:

\[
\epsilon[e(1+i)] = \frac{1 + \epsilon[u'(\cdot)]}{[1 + \epsilon[u'(q_s)]]\frac{1}{\eta}e[\alpha(e)][1 + \epsilon[u'(\cdot)] - [1 + \epsilon[\psi'(e)][\epsilon'[q_s] - \epsilon[u'(\cdot)]]}
\]

(6)
Example 3. Suppose \( \alpha(e) = e/(1 + e) \), \( u(q) = q^{1-\varepsilon}/(1 - \varepsilon) \) for \( 0 < \varepsilon < 1 \), \( c(q) = cq \) for \( c > 0 \), \( \psi(e) = \psi e \) for \( \psi > 0 \). Then the sign of \( \psi e(1 + i) \) is given by the sign of \( (1 - \varepsilon - \varepsilon(\eta - 1)/(1 + e) \) \( (\eta - 1)/(1 + e) \) This is clearly negative for \( e \) sufficiently high. Note that at \( e = 0 \), the relevant sign is that of \( (1 - \varepsilon) - e(\eta - 1) \). In particular if \( \eta \varepsilon > 1 \), then effort falls at a higher nominal interest rate.

Example 4. Let \( c(q) = q,u(q) = \overline{q}^{-\varepsilon}/(1 - \varepsilon) \), \( \psi(e) = e^2/2 \), \( \alpha(e) = e/(1 + e) \). Then the two curves can be characterized as

\[
q_s = \frac{\alpha(e) \frac{1}{(\eta-1)}}{\left( \frac{\eta}{\eta-1} \right)^{1/\varepsilon} (1 + i)^{1/\varepsilon} \mu} \\
e^2 = \frac{\eta}{(\eta - 1)^2} \mu q_s (1 + i)
\]

\( e \) cannot be solved in closed form. Instead, we obtain

\[
[e^\varepsilon(2\eta-1)-1(1 + e)^{1-\varepsilon}]^{-1/\varepsilon} (1 + i)^{(1-\varepsilon)/\varepsilon} = \left( \frac{\eta}{\eta-1} \right)^{2\varepsilon-1}
\]

from which it is clear that \( e \) is decreasing with \( i \) if \( \eta \varepsilon > 1 \). I have

\[
q = \left[ \frac{\eta}{\eta-1} (1 + i)(e/(1 + e))^{(\varepsilon-1)/(\eta-1)} \right]^{-1/\varepsilon} \\
\overline{q} = \left[ \frac{\eta}{\eta-1} (1 + i) \right]^{-1/\varepsilon} (e/(1 + e))^{1/(\varepsilon-1)}
\]

Hence, for general CRRA preferences, \( q \) is decreasing in \( e \) as long as \( \eta \varepsilon > 1 \). \( \overline{q} \) is always increasing in \( e \).

Example 5. I illustrate a sample equilibrium, choosing \( \eta = 6 \) and \( \varepsilon = 0.5 \). I consider \( i = 0.0 \) and \( i = 0.13 \) to highlight the movement from the Friedman rule to 10% inflation (\( \rho = 0.03 \)). Let the functional forms be as in Example 4. Then the equilibrium conditions are depicted in A.11:

Figure A.11: Equilibrium

\[
\text{Null}
\]

With \( i = 0.03 \), the equilibrium \((e, q_s)\) is given by \((0.3572, 0.5317)\), and with \( i = 0.13 \), it is given by \((0.3344, 0.4123)\). The decline in \( q_s \) is relatively sharper than the decline in \( e \).

Example 5 shows that a large movement in the nominal interest rate changes effort by less than 10%.

Appendix A.1.3. Social optimum

I write the (constrained) social planning problem. From an analogous argument to the basic model, symmetry simplifies the social planning problem into a choice of \( e \) and \( q_s \) so as to maximize the following welfare function:

\[
W(e, q_s) = u[\alpha(e) \frac{1}{(\eta-1)} \mu q_s] - \mu c(q_s) - \psi(e)
\]

The first order conditions are given by

\[
\alpha(e)^{1/(\eta-1)} u'(\overline{q}) = e'(q_s)
\]

\[
\mu q_s \alpha^{(2-\eta)/(\eta-1)}(\eta)(e) u' \overline{q} = (\eta - 1)\psi'(e)
\]
Equation (A.12) says that marginal utility equals aggregate marginal costs. Equation (A.13) says that the marginal cost of effort equals the marginal cost of the induced production. Taking the ratio of (A.13) and (A.12) and rearranging, we obtain
\[
\mu q_s c'(q_s) \epsilon[\alpha(e)] = (\eta - 1) e \psi'(e)
\] (A.14)

Equation (A.14) characterizes the optimal matching elasticity in terms of the marginal cost of effort, the marginal cost of production, and taste for variety scaled by the relative measure of sellers to buyers. Comparing (A.14) with (A.5), we note that implementing the social optimum would require
\[
\epsilon[\alpha(e)] = \frac{\eta}{\eta - 1}(1 + i)
\]
which is impossible, as the elasticity of the matching function is bounded above by 1. Hence, we have the following proposition.

**Proposition 11.** Every equilibrium is inefficient.

As the model with entry, if a deviation from the Friedman rule raises equilibrium welfare, it must do so via reducing effort.

**Appendix A.2. Derivation of marginal markups**

With generalized Nash bargaining, real balances can be expressed as \( z_\theta(q) = [1 - \Theta(q)]u(q) + \Theta(q)c(q) \), where \( \Theta(q) = \frac{\theta u'(q)}{\theta u'(q) + (1 - \theta)c'(q)} \) is the buyer’s share of the match surplus. The optimal consumption problem of the buyer in the DM is
\[
\max_{q \in [0, q^*]} \{ -i z_\theta(q) + \sigma[u(q) - z_\theta(q)] \}
\] (A.15)
which has solution
\[
\frac{u'(q)}{z_\theta'(q)} = 1 + \frac{i}{\sigma}
\] (A.16)
Substituting \( z_\theta'(q) \), dividing both sides by \( c'(q) \), and rearranging yields the marginal markup.

Under proportional bargaining, as long as buyer spends all real balances in the DM, then \( z_\theta(q) = \theta c(q) + (1 - \theta)u(q) \).

The optimal consumption problem of the buyer in the DM can be written as
\[
\max_{q \in [0, q^*]} \{ -i z_\theta(q) + \sigma \theta[u(q) - c(q)] \}
\] (A.17)
This problem is concave provided \( \frac{\theta}{1 - \theta} > \frac{i}{\sigma} \), and has solution
\[
\frac{u'(q) - c'(q)}{z_\theta'(q)} = \frac{i}{\theta \sigma}
\] (A.18)
Substituting \( z_\theta'(q) \), dividing through by \( c'(q) \), and rearranging algebraically yields the marginal markup.

**Appendix A.3. Second order condition for multiple equilibria with entry**

The second derivative condition of the buyer’s objective function is
\[
u''(\overline{q}) \alpha(\mu) \frac{1}{\overline{q} - q} + u'(\overline{q}) \frac{1}{\eta} \frac{q(j) \alpha(\mu) \frac{n}{q(j)^2} - \overline{q}}{q(j)^2} \leq 0
\] (A.19)
at equilibrium, which simplifies to
\[ \epsilon[u'(\eta)] + \frac{1}{\eta} \alpha(\mu) \frac{1}{1-\epsilon} - \alpha(\mu) \frac{\eta}{1-\epsilon} \leq 0 \] (A.20)

In the case of CRRA preferences, (A.20) reduces to
\[ \alpha(\mu) \frac{1}{1-\epsilon} - \alpha(\mu) \frac{\eta}{1-\epsilon} \leq \eta \epsilon \] (A.21)

For \( \alpha(\mu) = \frac{\mu}{1+\mu}, \eta = 3, \epsilon = 0.25 \), it is easy to check that the second order condition holds for any \( \mu \), as indicated in Figure A.12. There we plot \( h(\mu) = \alpha(\mu) \frac{1}{1-\epsilon} - \alpha(\mu) \frac{\eta}{1-\epsilon} - \eta \epsilon \). Note that without loss of generality we can restrict attention to \( \mu \) on \([0, 1]\), as \( h(\mu) \) is always negative for \( \mu > 1 \).

Figure A.12: Second order condition

Appendix A.4. Money demand functions

Appendix A.4.1. Basic model

Money demand is defined as the ratio of aggregate money balances to aggregate nominal output. For this model, this can be written as
\[ L = \frac{\sigma z}{\sigma z + x} \] (A.22)

where \( x \) is the quantity of the CM good. I make the restriction \( \sigma = 1 \) and \( \mu = 1 \). I obtain \( z = \frac{\eta}{\eta-1} c(q_s) = \frac{\eta}{\eta-1} q_s \), given linearity of cost. Using \( q_s = \left[ \frac{\alpha(1-\epsilon)/(\eta-1)}{\eta/(1+i)} \right]^{1/\epsilon} \), we obtain
\[ L = \left[ 1 + A \left[ \frac{\left( \frac{\eta}{\eta-1} \right)^{1-\epsilon} (1+i)}{\alpha(1-\epsilon)/(\eta-1)} \right]^{1/\epsilon} \right]^{-1} \] (A.23)

Appendix A.4.2. Entry of sellers with fixed effort

Real balances are given by \( z(q_s) = \frac{\eta}{\eta-1} c(q_s) q_s \). We use \( c(q_s) = q_s \) and \( u(q) = q^{1-\epsilon}/(1 - \epsilon) \). From the zero profit condition, \( q_s = (\eta - 1)k \). Moreover, \( \mu \) is implicitly characterized by \( (\mu q_s)^{\epsilon} = \frac{\alpha(1-\epsilon)/(\eta-1)}{\eta/(1+i)} \). Here there is no explicit solution for money demand.

Appendix A.4.3. Additively separable preferences

Real balances in the economy are given by \( z(q_s) = \mu q_s c'(q_s) \). Taking \( c(q_s) = q_s \), money demand is given by \( L = \frac{\epsilon}{\epsilon + A} \left[ 1 + \frac{A(1 - r_u(q))}{\mu q_s c'(q_s)} \right]^{-1} \). We adopt the utility function \( u(q) = q^{1-\epsilon}/(1 - \epsilon) + B q \), so that \( r_u(q) = \frac{\epsilon}{1 + B q} \).
Appendix A.4.4. Variable search intensity

I obtain a closed form solution and obtain the money demand by restricting \( \psi(e) = \psi^2/2 \) and otherwise using the same specification as in the basic case. I obtain a similar expression for \( q_s \) as in the basic model, except now as a function of \( e \):

\[
q_s = \left[ \frac{\alpha(e)(1-\varepsilon)/(\eta-1)}{\eta-1(1+i)} \right]^{1/\varepsilon}
\]

where \( e = \sqrt{\frac{\eta}{(\eta-1)^2}} q_s (1+i) \). Letting \( e^* \) denote the equilibrium value of search effort, the money demand is given by

\[
L = \left\{ 1 + A \left[ \frac{\left( \frac{\eta}{\eta-1} \right)^{1-\varepsilon} (1+i)^{1-\varepsilon}}{\alpha(e^*)(1-\varepsilon)/(\eta-1)} \right]^{1/\varepsilon} \right\}^{-1} \tag{A.24}
\]

Appendix A.5. Effect of a change in the elasticity of the matching function on welfare costs of inflation

Results pending

Table A.6: Welfare costs of inflation under different elasticities of matching function

<table>
<thead>
<tr>
<th>( \mu_p )</th>
<th>( \Delta )</th>
<th>( \varepsilon )</th>
<th>( A )</th>
<th>( k )</th>
<th>( \Delta )</th>
<th>( \varepsilon )</th>
<th>( A )</th>
<th>( k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.01</td>
<td>0.0184</td>
<td>0.0661</td>
<td>0.9111</td>
<td>0.0103</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.05</td>
<td>0.0227</td>
<td>0.0748</td>
<td>0.4776</td>
<td>0.0206</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.1</td>
<td>0.0325</td>
<td>0.1510</td>
<td>0.2351</td>
<td>0.0693</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.2</td>
<td>0.0519</td>
<td>0.2024</td>
<td>0.1141</td>
<td>0.0634</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.3</td>
<td>0.0823</td>
<td>0.2268</td>
<td>0.0741</td>
<td>0.0357</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Appendix A.5.1. Firm diversification

We generalize the free entry model with endogenous choice of varieties per firm. This is motivated by the finding in Bernard, Redding, and Schott (2010) that 94% of new products in manufacturing occur within existing facilities.

Following Dong (2010), firm \( f \) produces a unique set of goods \( \varphi_f \), where \( \varphi_f \cap \varphi_f' = \emptyset \) for all \( f \) and \( f' \). In the CM, firms choose whether to enter and the capacity of goods \( N \) in which to invest. Hence, entry and capacity costs are given by \( k(N) \), where \( k(0) = \overline{k}, k' > 0, k'' > 0 \) and \( k''(0) = 0 \). \( \overline{k} \) are costs associated with setting up the firm independent of investment in capacity for production of varieties. It serves an important technical purpose for existence of equilibrium, as we shall see. The quantity index for buyers is defined as

\[
\bar{q} = \int_0^{\alpha(\mu)} \left( \int_0^N q_j^{\eta-1} \, dn \right) dj^{\frac{\eta}{\eta-1}} \tag{A.25}
\]

This is only with minor loss of generality. Firms would never wish to produce identical products if they have the option of differentiating their products at the same cost. Hence, firms producing identical products is only relevant to the extent that doing so is less costly than differentiating their product. Also, if two firms produce the same product, then results are sensitive to the nature of competition (Bertrand vs. Cournot) and whether consumers access both firms or not.
where \( q_{j,n} \) is the quantity of variety \( n \) offered by firm \( j \). The problem of the buyer yields

\[
(1 + i)p_{j,n} = [N\alpha(\mu)]^{\frac{1}{\eta-1}} u'(\bar{q})(\frac{\bar{q}}{q_{j,n}})^{\frac{1}{\eta}}
\]  

(A.26)

Anticipating product symmetry, let \( q = q_{n,j} \) for all \( n, j \), so that \( \bar{q} = [N\alpha(\mu)]^{\frac{1}{\eta-1}} q \) and the quantity of each variety is given by \( q_s = \frac{\alpha(\mu)}{\mu} q \). This implies \( \bar{q} = [N\alpha(\mu)]^{\frac{1}{\eta-1}} \mu N q_s \). \( \bar{q} \) thus consists of total production \( \mu N q_s \) scaled according to the number of varieties each consumer accesses and the taste for variety \( [N\alpha(\mu)]^{\frac{1}{\eta-1}} \).

The marginal markup in equilibrium is given by

\[
\frac{u'((N\alpha(\mu))^{\frac{1}{\eta-1}} \mu N q_s)}{c'(q_s)} = \frac{\eta}{\eta-1} (1 + i)
\]  

(A.27)

The maximization problem over the number of varieties is given by

\[
\max_N \left\{ -k(N) + N \left[ \frac{\eta}{\eta-1} c'(q_s) q_s - c(q_s) \right] \right\}
\]  

(A.28)

where \( q_s = \alpha_s q \), given the inverse demand of the buyer \( q_j \). Given the assumptions on \( k(\cdot) \), there is an interior solution given by

\[
k'(N) = \frac{\eta}{\eta-1} c'(q_s) q_s - c(q_s)
\]  

(A.29)

The firm equates the marginal cost of a new production line with the profits of the associated variety. In contrast to Dong, the benefit consists of selling greater varieties to the same set of customers rather than increasing the probability of supplying consumers’ desired variety.

Firm size is given by \( N q_s \), the measure of varieties multiplied by the quantity of each variety. The free entry condition, in turn, is characterized by

\[
k(N) + N c(q_s) = N \frac{\eta}{\eta-1} c'(q_s) q_s
\]  

(A.30)

Noting that average costs are given by \( \frac{k(N) + N c(q_s)}{N q_s} \), so that the ratio of average to marginal costs is given by \( \frac{k(N) + N c(q_s)}{q_s c'(q_s)} = \frac{\eta}{\eta-1} \), which is the standard free entry condition with \( k(N)/N \) in place of \( k \).

Note that (A.29) and (A.30) can be combined to yield

\[
k'(N) = \frac{k(N)}{N}
\]  

(A.31)

This result is a consequence of free entry, and is independent of CES. In particular, firm entry drives average costs of investment in capacity to its minimum, which occurs at \( k'(N) = \frac{k(N)}{N} \), or \( \varepsilon_k(N) = 1 \). Since \( k(\cdot) \) is convex, a solution to (A.31) requires \( \overline{k} > 0 \). Note that, with \( N^c \) determined from (A.31), \( q_s \) can be determined from (A.30) and \( \mu \) can be determined from (A.27). Hence, equilibrium is a triple \((N, q_s, \mu)\) that satisfies (A.31), (A.30), and (A.30). Due to the recursive structure, the conditions for existence and uniqueness are identical to the free entry equilibrium. Analogously to the free entry model, define \( \lambda(\mu) = [N^c\alpha(\mu)]^{\frac{1}{\eta-1}} u'((N^c\alpha(\mu))^{\frac{1}{\eta-1}} \mu N^c q^*_s) \), which is marginal utility as a function of the sellers, given \( N = N^c \) and \( q_s = q^*_s \).

From the recursive structure of equilibrium, \( N \) and \( q_s \) do not depend on \( i \). A higher nominal interest rate shifts the cost of holding real balances higher, and through (A.27) reduces the entry of sellers, provided \( \lambda'(\mu) < 0 \). Hence, inflation reduces variety only by affecting the entry of sellers, not by reducing the varieties of each firm. It is easy to see that in the absence of free entry, however, inflation would reduce the varieties
offered by each firm. This arises from the basic model, in which \( q_s \) decreases inflation. Here, \( N \) would be decreasing with \( q_s \) and, hence, with inflation. Thus, with a fixed measure of sellers, we recover the result in Dong that both quantity and variety decrease with inflation.

This generalization provides an account of the tradeoff between firm diversification vs. specialization. Producing more varieties allows a firm to extract additional profits given by \( \frac{n}{n-1} c'(q_s) q_s - c(q_s) \), but faces convex costs of investment in capacity.


URL http://ideas.repec.org/p/nbr/nberwo/13646.html


