University of Hawai`i at Mānoa Department of Economics Working Paper Series

Saunders Hall 542, 2424 Maile Way,
Honolulu, HI 96822
Phone: (808) 956 -8496
www.economics.hawaii.edu

Working Paper No. 14-29R

Endogenous Search, Price Dispersion, and Welfare

By

Liang Wang

October 2014

*Revised Version of WP No. 11-13

Endogenous Search, Price Dispersion, and Welfare*

Liang Wang†
University of Hawaii Manoa

September 18, 2014

Abstract

This paper studies the welfare cost of inflation in a frictional monetary economy with endogenous price dispersion, which is generated by sellers posting prices and buyers costly searching for low prices. We identify three channels through which inflation affects welfare. The interaction of real balance channel and price posting channel generates a welfare cost, at 10% annual inflation, equal to 3.23% of steady state consumption; if either channel is shut down, the welfare cost decreases to less than 0.15%. Search channel reduces welfare cost by more than 50%. The aggregate effect of inflation on welfare is nonmonotonic.

Keywords: Inflation, Price Dispersion, Search, Welfare

JEL: E31, E40, E50, D83

*I am thankful to Kenneth Burdett, Guido Menzio, and in particular Randall Wright for invaluable guidance and support. Additional thanks go to Aleksander Berentsen, Richard Dutu, Allen Head, David Parsley, Adrian Peralta-Alva, Christopher Waller, and participants at the Search and Matching Workshop, AEA Annual Meeting, Midwest Macro Meetings, Econometric Society NASM, the seminars at the University of Pennsylvania and the Federal Reserve Bank of St. Louis. I am grateful to Guillaume Rocheteau, Richard Dutu, and David Parsley for sharing their data. All errors are mine.

†Department of Economics, 2424 Maile Way, Honolulu, HI 96822. Email: lwang2@hawaii.edu.

1 Introduction

There is a long tradition of thinking that the welfare cost of inflation has two major sources: the opportunity cost of holding money and price dispersion. In this paper, we revisit the classical question on welfare cost in a general equilibrium monetary model with price dispersion. We build on the monetary search framework of Lagos and Wright (2005), and integrate the analysis of endogenous price dispersion in a frictional goods market by Burdett and Judd (1983).

In this economy, buyers want to carry money despite a positive opportunity cost because they can use it as a medium of exchange in the bilateral trade market with frictions, in which other means of payment are infeasible due to anonymity and imperfect monitoring technology. Search frictions not only make money essential, but also grant sellers monopolistic power to set prices. In the frictional market, buyers can only observe a subset of prices, and they need to pay a cost to acquire more price information. Therefore, when a seller serves buyers who do not search for lower prices, he is able to sell at prices higher than marginal cost, but then the seller can also post a lower price to attract buyers who search. This trade-off faced by sellers generates endogenous price dispersion.

We identify three channels through which inflation affects welfare: real balance, price posting, and endogenous search. When inflation increases, holding money becomes more costly, and buyers reduce their real balances. As a result, consumption decreases, and so does welfare. This is the real balance channel. At the same time, buyers also want to search harder, in order to find a lower price and purchase more goods. So search channel increases consumption and welfare. Price posting channel does not respond to inflation directly, but changes with the other two channels. Sellers certainly post different prices if buyers bring less money to trade or search harder for low prices, and price changes affect consumption and welfare subsequently.

In order to understand the aggregate effect of inflation on welfare, as well as the effects through each individual channel, we calibrate the model to match U.S. data on money demand and price dispersion, measure the welfare cost of inflation following Lucas (2000), and decompose the aggregate effect of inflation by shutting down each channel separately. Taking the economy with zero inflation as a benchmark, we find the welfare cost of 10% annual inflation is worth 3.23% of consumption. However, if either real balance or price posting channel is shut down, welfare cost sharply decreases to less than 0.15% of consumption.

The decomposition exercise shows that the large welfare cost of inflation is caused by the interaction of real balance and price posting channel. If inflation increases, real balance and consumption decrease. Since buyer's expenditure is not sensitive to price changes, sellers respond by posting even higher prices, which further reduces consumption and welfare. Search channel has two effects on welfare. First, if buyers search more, they are more likely to trade at lower prices, and hence consumption increases. Second, search activities increase competition among sellers, and drive down overall real price level. Our quantitative exercise suggests that the second effect of search channel on welfare is more important, and it reduces the welfare cost of inflation by 50%.

The model predicts a nonmonotonic effect of inflation on welfare. Even at the Friedman rule, sellers still post prices above marginal cost. A small deviation from the Friedman rule can improve welfare by encouraging buyers to search harder for low prices and increasing competition among sellers. However, as the cost of holding money increases more, real balance drops quickly, and it has a first-order effect on consumption. The joint negative effect of real balance and price posting channel becomes larger than the positive effect of search channel, and welfare decreases with inflation.

This paper fits into a long literature since Bailey (1956), which studies the welfare cost of inflation through real balance channel. Among recent papers, Cooley and Hansen (1989) use a cash-in-advance constraint to introduce money into a real business cycle and find the welfare cost of 10% inflation equals to 0.52% of steady state consumption. Lucas (2000) surveys research on welfare cost in different frameworks, and estimates the welfare cost of 10% inflation to be less than 1% of real income. We find a larger welfare cost, but the effect of inflation through real balance channel alone is in the same range as previous findings.

Following the New Monetarist literature, in our model, money is introduced via search frictions. Many papers in this literature focus on real balance channel and find large welfare cost of inflation due to the holdup problem associated with bargaining, such as Lagos and Wright (2005), Craig and Rocheteau (2008a), and Rocheteau and Wright (2009). Because of endogenous search channel, the welfare cost of inflation in our model is slightly smaller than the findings in those papers. However, we consider price posting instead of bargaining, and present a new source of large welfare loss, even without the hold up problem.

Several papers in this literature generate endogenous price dispersion in equilibrium. Head et al. (2012) also integrate the Lagos-Wright framework with Burdett and Judd (1983), but search intensity is exogenous. They show that the model can match the empirical evidence of price changes very well. In the same framework with indivisible goods, Liu et al. (2014) study money and credit as alternative means of payment. These two papers do not examine welfare. Dutu, Julien, and King (2012) introduce second-price auction into the Lagos-Wright framework, and study the welfare cost (or gain) of price dispersion. In another paper by Head and Kumar (2005), they build on the large household framework in Shi (1997), and study the relationship between inflation and price dispersion. They also discuss welfare and present similar qualitative findings as in our paper, but no quantitative results since they do not do calibration. In a cashless search model with price dispersion, Benabou (1988, 1992) and Diamond (1993) study the theoretical connection between inflation and efficiency. They ignore the real balance channel, which greatly affects welfare.

There are other frictions which can also generate price dispersion, such as Calvo pricing and menu costs in the New Keynesian literature. Burstein and Hellwig (2008) study a variety of New Keynesian models, and calculate a welfare cost, only under Calvo pricing, similar to the magnitude in Lagos and Wright (2005). Craig and Rocheteau (2008b) combine the Lagos-Wright framework with menu costs and find the optimal inflation rate is away from the Friedman rule. Aruoba and Schorfheide (2011) combine Lagos-Wright with Calvo pricing and find that the welfare distortions created by search frictions are of similar magnitude as the distortions created by the New Keynesian friction. All three papers include real balance and price posting channels. We also study endogenous search channel, and focus on the welfare implication of the interaction of different channels.

The remainder of the paper is organized as follows. Section 2 lays out the environment of the model, and we solve for monetary equilibrium in Section 3. Then, Section 4 presents quantitative analysis, including calibration exercise and welfare analysis. Section 5 concludes the paper. Additional technical details and proofs are in the Appendix.

2 The Environment

Time is discrete. Each period is divided into two subperiods. In the first subperiod, there is a decentralized market (hereafter DM) and goods are traded bilaterally. In the second subperiod, the market is centralized (hereafter CM) and there is Walrasian trade in the market. A continuum of buyers and sellers, each with measure one, live forever. Following Rocheteau and Wright (2005), assume both types produce and consume in the CM, but they act differently in the DM. Buyers want to consume but cannot produce, while sellers can produce but do not want to consume. All economic agents are assumed to be anonymous in the DM, and there is imperfect monitoring or record keeping technology. These assumptions, as well as the lack of double coincidence of wants, make a medium of exchange, which is called money, essential.[1] Money is storable and perfectly divisible.

M_t denotes money supply in period t, and it grows according to $M_{t+1} = (1+\gamma)M_t$, where M_{t+1} is money supply in the next period $t+1$. New money is injected by lump-sum transfers, or withdrawn by lump-sum taxes if $\gamma < 0$, at the beginning of the CM. For simplicity, we assume that transfer or tax goes equally to each buyer.[2]

In period t, the buyer's instantaneous utility function is

$$U_t^b(x_t, h_t, q_t) = u(q_t) + v(x_t) - h_t,$$

where q_t is the quantity of the DM goods consumed, x_t is the quantity of the CM goods consumed, and h_t is the quantity produced. The CM goods are produced one-for-one using labor. The lifetime utility of a buyer is $\sum_{t=0}^{\infty} \beta^t U_t^b$. Assume that $u(q)$ has the CRRA form with risk aversion coefficient $\sigma < 1$ and $u(0) = 0$, $u'(q) > 0$, $u''(q) < 0$ for all q. Also assume $v'(x) > 0$ and $v''(x) < 0$ for all x, and there exists $x^* > 0$ such that $v'(x^*) = 1$. Similarly, the instantaneous utility of a seller is

$$U_t^s(x_t, h_t, q_t) = -cq_t + v(x_t) - h_t,$$

[1] For more discussions on the essentiality of money, please refer to a recent paper by Wallace (2011).

[2] Alternatively, we can assume that transfer or tax goes equally to each agent or each seller, and equilibrium results do not change.

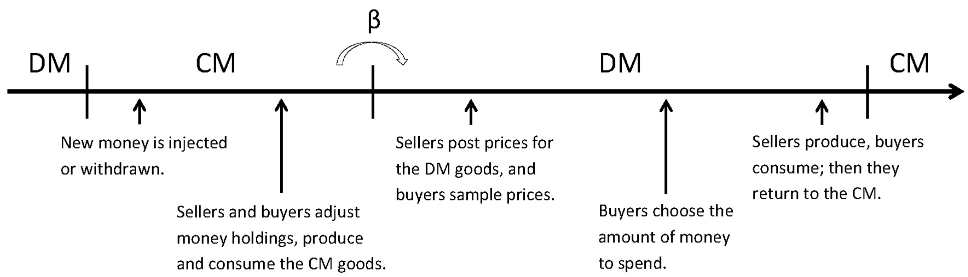

Figure 1: Timeline

where q_t, x_t, and h_t have the same definitions as in the buyer's utility function.[3] The lifetime utility of a seller is $\sum_{t=0}^{\infty} \beta^t U_t^s$. Let $u'(q) = c$ hold for some $q^* > 0$.

In the CM, the price of consumption good x is normalized to one. The relative price of money in terms of x is ϕ_t in period t and the price of x_t in terms of money is $1/\phi_t$. Then, the CM consumption good becomes the numeraire in the economy. β is the discounting factor between today's DM and tomorrow's CM. This paper focuses on the case in which $\beta < 1+\gamma$. Assume inflation is forecasted perfectly and the Fisher equation holds. Hence, the nominal interest rate i is equal to $(1+\gamma-\beta)/\beta$, and $\beta < 1+\gamma$ implies $i > 0$.

3 Search and Price Dispersion

This paper studies a market structure in which sellers post prices and buyers know the price distribution but cannot observe all the prices. Burdett and Judd (1983) study a similar search protocol in a non-monetary model of indivisible goods. In this random search environment, buyers have the freedom to sample one or two prices, or equivalently, to visit one or two sellers with a search cost. They have knowledge about the price distribution but not about an individual price or an individual seller. Hence, a buyer cannot direct his search to the seller with the lowest price, and he has to visit a seller without knowing his price ex ante.

Figure 1 presents the timeline of the events. At the beginning of the CM in each period, new money is injected or withdrawn by the government. Then, both sellers and buyers adjust their monetary balances, produce, and consume the CM goods. After agents enter the DM in the next period, each seller posts prices for the DM goods, and he is committed

[3] To simplify the analysis of the model, we assume that the seller's marginal cost of production is constant. The intuition behind does not change with more general forms of the cost function.

to producing and selling any quantity of the goods at the price posted. Every buyer then chooses his search intensity by sampling one or two prices from the price distribution of the DM goods.[4] After that, he still needs to decide how much money to spend in a trade. Finally, each buyer trades with one seller. The seller produces, the buyer consumes, and then they return to the CM.

Sellers can randomize over a range of prices in the price posting stage. The induced price distribution in the DM in period t is denoted as F_t with support $\mathcal{Z}_{F_t} = \left[\underline{p}_t, \bar{p}_t\right]$. Based on the knowledge about F_t, buyers make their decisions on search intensity. After entering the DM, every buyer can observe one price for free and he can choose to pay cost k to observe a second price. Buyers who choose not to sample a second price are type-1 buyers, and the others are type-2 buyers. The measure of type-2 buyers is denoted by $\alpha_t \in [0, 1]$, and $1 - \alpha_t$ is the measure of type-1 buyers.

The real balance that an agent carries in period t is denoted as $z_t = m_t \phi_t$. Starting from this point, our analysis focuses on stationary monetary equilibrium where aggregate real variables stay constant. This implies that $\phi_t M_t = \phi_{t+1} M_{t+1}$ and $\phi_t / \phi_{t+1} = 1 + \gamma$. The rate of nominal price change, i.e., the inflation or deflation rate in the CM is equal to the money growth rate $1 + \gamma$. In the DM, the rate of nominal price change is slightly more complicated due to the existence of price dispersion, and it may be greater or smaller than $1 + \gamma$. From this point onwards we suppress the time subscript and use $\hat{}$ to denote the variables of the next period. $W^b(z)$ and $V^b(z)$ are buyer's value functions in the CM and DM, respectively, and $W^s(z)$ and $V^s(z)$ are seller's value functions. We proceed first with the buyer's optimization problem.

3.1 Buyer's Optimization Problem

When a buyer enters the CM, he needs to decide whether he is going to observe one or two prices in the following DM. This choice affects the buyer's decision on production, consumption, and money holding in the CM. A type-1 buyer faces the following optimization

[4] To ease the presentation, we limit the maximum number of prices that a buyer can sample to two. One can extend the logic of Claim 1 in Burdett and Judd (1983, p. 962) and prove that in equilibrium with price dispersion, buyers endogenously choose not to sample more than two prices.

problem in the recursive form.

$$W_1^b(z) = \max_{x_1, h_1, \hat{z}_1} \left\{ v(x_1) - h_1 + \beta V_1^b(\hat{z}_1) \right\} \tag{1}$$
$$s.t.\ h_1 + z + T = x_1 + (1+\gamma)\hat{z}_1$$

$W_1^b(z)$ represents the value function in the CM and $V_1^b(\hat{z}_1)$ is the value of observing one price in the DM. The buyer produces the CM goods using labor as input, consumes, and adjusts his real balance of the next period $\hat{z}_1$. $T = \gamma\phi M$ is the transfer payment made by the government.

A type-2 buyer chooses to observe a second price with cost k in the DM, and his optimization problem in the CM is characterized as

$$W_2^b(z) = \max_{x_2, h_2, \hat{z}_2} \left\{ v(x_2) - h_2 + \beta V_2^b(\hat{z}_2) \right\} \tag{2}$$
$$s.t.\ h_2 + z + T = x_2 + (1+\gamma)\hat{z}_2$$

We define V_1^b and V_2^b explicitly later. Therefore, at the beginning of the CM, a buyer chooses the option with the higher value.

$$W^b(z) = \max \left\{ W_1^b(z), W_2^b(z) \right\}$$

Let us look at $W_1^b(z)$ first. Substitute h in the value function by the budget constraint, and (1) becomes

$$W_1^b(z) = z + W_1^b(0) \tag{3}$$

where $W_1^b(0) = \max_{x, \hat{z}_1} \left[v(x) - x + T - (1+\gamma)\hat{z}_1 + \beta V_1^b(\hat{z}_1) \right]$. The buyer's optimal decision of $\hat{z}_1$ does not depend on his current money holding. Similarly, we have

$$W_2^b(z) = z + W_2^b(0) \tag{4}$$

and $W_2^b(0) = \max_{x, \hat{z}_2} \left[v(x) - x + T - (1+\gamma)\hat{z}_2 + \beta V_2^b(\hat{z}_2) \right]$. Therefore,

$$W^b(z) = z + \max \left\{ W_1^b(0), W_2^b(0) \right\} = z + W^b(0) \tag{5}$$

This convenient result is due to the assumption of quasi-linear utility function in the CM, which yields a degenerate distribution of buyers' money holdings in the DM.

Now we turn to the DM. The value function of a type-1 buyer $V_1^b(z_1)$ is given by

$$V_1^b(z_1) = \int_{\underline{p}}^{\bar{p}} \left\{ u\left(\frac{d^*(p; z_1)}{p}\right) + W^b\left[z_1 - d^*(p; z_1)\right] \right\} \mathrm{d}F(p), \tag{6}$$

where $d^*(p; z_1)$ represents the buyer's optimal expenditure on the DM goods, which depends on his money holding z_1 and transaction price p. After the buyer pays $d^*(p; z_1)$ for the DM goods, he still carries a real balance of $z_1 - d^*(p; z_1)$, and $W^b[z_1 - d^*(p; z_1)]$ represents the continuation value of entering the next CM.

A type-2 buyer faces the following value function

$$V_2^b(z_2) = \int_{\underline{p}}^{\bar{p}} \left\{ u\left(\frac{d^*(p; z_2)}{p}\right) + W^b\left[z_2 - d^*(p; z_2)\right] \right\} \mathrm{d}\left[1 - (1 - F(p))^2\right] - k, \tag{7}$$

where $d^*(p; z_2)$ is defined similarly as $d^*(p; z_1)$, and $1 - (1 - F(p))^2$ is the distribution of transaction price, i.e., the lower of two observed prices.

In order to solve for $d^*(p; z)$, the buyer's optimal expenditure function, we apply the linearity of $W^b(z)$ from (5) and rewrite (6) and (7) as

$$V_1^b(z_1) = \int_{\underline{p}}^{\bar{p}} \left\{ u\left(\frac{d^*(p; z_1)}{p}\right) - d^*(p; z_1) \right\} \mathrm{d}F(p) + W^b(z_1) \tag{8}$$

and

$$V_2^b(z_2) = \int_{\underline{p}}^{\bar{p}} \left\{ u\left(\frac{d^*(p; z_2)}{p}\right) - d^*(p; z_2) \right\} \mathrm{d}\left[1 - (1 - F(p))^2\right] - k + W^b(z_2). \tag{9}$$

It is obvious that $d^*(p; z)$ is the solution to the following problem.

$$\max_{d \geq 0} u\left(\frac{d}{p}\right) - d$$
$$s.t.\ d \leq z$$

A buyer chooses how much money to spend on the DM goods, and he cannot spend more than what he carries. The following Lemma explicitly characterizes $d^*(p; z)$.

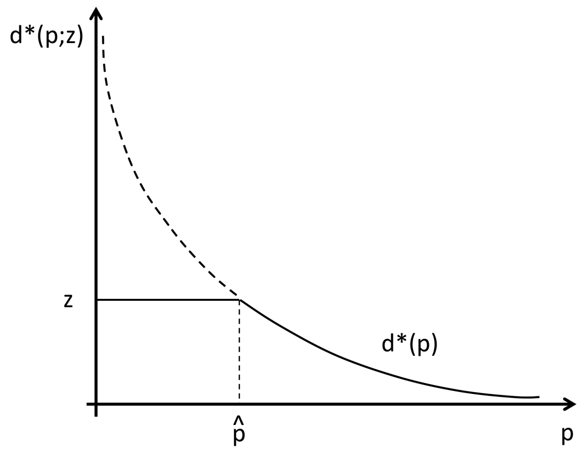

Figure 2: Buyer's Optimal Spending Rule in the DM

Lemma 1 *The buyer's optimal spending rule is given by*

$$d^*(p;z) = \begin{cases} z, & \text{if } p < \hat{p} \\ d^*(p), & \text{otherwise} \end{cases}$$

where $\hat{p}$ and $d^*(p)$ satisfy $u'(z/\hat{p}) = \hat{p}$ and $u'(d^*(p)/p) = p$, respectively, and $\partial\hat{p}/\partial z < 0$, $\partial d^*(p)/\partial p < 0$.

Proof. see Appendix. ■

The risk aversion coefficient σ characterizes the buyer's price elasticity of demand. When σ is less than one, the buyer's price elasticity of demand is greater than one, and the expenditure elasticity is less than one. Then, his expenditure on the DM goods $d^*(p;z)$ decreases as he faces a higher price level. A buyer cannot spend more than his monetary constraint at low price levels even though he desires to do so. $\hat{p}$ is the cutoff price level at which the buyer's monetary constraint starts to relax, and he spends less than the total money holding when the price is higher than $\hat{p}$. This situation is illustrated in Figure 2.

We substitute (8) and (9) into (3) and (4), and the buyer's Bellman's equations in the CM become

$$\begin{aligned} W_1^b(z) &= z + \max_{x,\hat{z}_1} \Big\{ v(x) - x + T - (1+\gamma)\hat{z}_1 + \beta W^b(\hat{z}_1) \\ &\quad + \beta \int_{\underline{p}}^{\bar{p}} \left[u\left(\frac{d^*(p;\hat{z}_1)}{p} \right) - d^*(p;\hat{z}_1) \right] \mathrm{d}F(p) \Big\} \end{aligned} \tag{10}$$

and

$$
\begin{aligned}
W_2^b(z) &= z + \max_{x,\hat{z}_2} \Big\{ v(x) - x + T - (1+\gamma)\hat{z}_2 + \beta W^b(\hat{z}_2) - \beta k \\
&\quad + \beta \int_{\underline{p}}^{\bar{p}} \left[u\left(\frac{d^*(p;\hat{z}_2)}{p}\right) - d^*(p;\hat{z}_2) \right] \mathrm{d}\left[1 - (1-F(p))^2\right] \Big\}
\end{aligned} \tag{11}
$$

The optimal decision of x satisfies $v'(x^*) = 1$, and it does not depend on $\hat{z}_1$ or $\hat{z}_2$.

According to Lemma (1), $d^*(p;z)$ has different expressions depending on the relationship between p and $\hat{p}$. Therefore, in order to characterize the buyer's optimal decision on $\hat{z}_1$ and $\hat{z}_2$, we first need to establish the following result characterizing the relationship of $\hat{p}$ and $\mathcal{Z}_F$.

Lemma 2 *In the optimization problem in (10) and (11), both type-1 and type-2 buyers always choose real balances $\hat{z}_1$ and $\hat{z}_2$ such that $\hat{p}_1 > \underline{p}$ and $\hat{p}_2 > \underline{p}$.*

Proof. see Appendix. ■

The intuition of Lemma 2 is straightforward. Consider type-1 buyers as an example. If the cutoff price $\hat{p}_1$ is smaller than the lower limit of price distribution, the buyer's real balance does not affect the surplus from trade in the DM since $d^*(p;z_1) = d^*(p)$. The marginal benefit of bringing more money to the DM is zero, while the marginal cost is still positive. Thus, a buyer wants to reduce his real balance until there is a positive marginal gain related to the action of carrying money, which only happens when $\hat{p}_1$ exceeds $\underline{p}$. This intuition is true for both type-1 and type-2 buyers.

We proceed to characterize the optimal decisions of buyers. Taking $F(p)$ as given, a type-1 buyer chooses $\hat{z}_1$ to solve the maximization problem in (10). We can rewrite the buyer's value function in the CM as the following

$$
\begin{aligned}
W_1^b(z) &\simeq \max_{\hat{z}_1} \Bigg\{ -(1+\gamma)\hat{z}_1 + \beta \int_{\underline{p}}^{\hat{p}_1} \left[u\left(\frac{\hat{z}_1}{p}\right) - \hat{z}_1 \right] \mathrm{d}F(p) \\
&\quad + \beta \int_{\hat{p}_1}^{\bar{p}} \left[u\left(\frac{d^*(p)}{p}\right) - d^*(p) \right] \mathrm{d}F(p) + \beta W^b(\hat{z}_1) \Bigg\},
\end{aligned} \tag{12}
$$

and the terms unrelated to $\hat{z}_1$ are omitted. So the buyer's optimal real balance $\hat{z}_1^*$ satisfies

$$
\int_{\underline{p}}^{\hat{p}_1} \left[u'\left(\frac{\hat{z}_1^*}{p}\right) \frac{1}{p} - 1 \right] \mathrm{d}F(p) = i. \tag{13}
$$

The buyer's marginal gain of holding money, which is the left hand side of (13), decreases as $\hat{z}_1^*$ increases. Holding everything else constant, there is less marginal gain as a buyer holds more money. From a partial equilibrium point of view, as the money growth rate increases, the nominal interest rate rises, the marginal cost of holding money gets bigger, and the buyer decides to carry a smaller real balance.

The price distribution $F(p)$ also affects the buyer's money holding in the following way. If $\tilde{F}(p)$ first-order stochastically dominates $F(p)$, a buyer carries less money with $\tilde{F}(p)$. Because he faces a market with a smaller probability of getting a low price, implying a smaller probability of becoming cash constrained and a smaller marginal gain of holding extra money. Therefore, the marginal benefit of carrying money is smaller than the marginal cost, and the buyer wants to reduce his money holding.

Next, a type-2 buyer chooses $\hat{z}_2$ to solve the maximization problem in (11) and we can rewrite the value function in a similar way as

$$\begin{aligned} W_2^b(z) \simeq \max_{\hat{z}_2} \Bigg\{ & -(1+\gamma)\hat{z}_2 + \beta \int_{\underline{p}}^{\hat{p}_2} \left[u\left(\frac{\hat{z}_2}{p}\right) - \hat{z}_2 \right] \mathrm{d}\left[1-(1-F(p))^2\right] \\ & + \beta \int_{\hat{p}_2}^{\bar{p}} \left[u\left(\frac{d^*(p)}{p}\right) - d^*(p) \right] \mathrm{d}\left[1-(1-F(p))^2\right] + \beta W^b(\hat{z}_2) \Bigg\}. \end{aligned} \tag{14}$$

Then, the buyer's optimal real balance $\hat{z}_2^*$ satisfies the following FOC

$$\int_{\underline{p}}^{\hat{p}_2} \left[u'\left(\frac{\hat{z}_2^*}{p}\right)\frac{1}{p} - 1 \right] \mathrm{d}\left[1-(1-F(p))^2\right] = i, \tag{15}$$

and the same intuition as $\hat{z}_1^*$ applies.

3.2 Seller's Optimization Problem

In the CM, the seller's value function is

$$\begin{aligned} W^s(z) &= \max_{x,h,\hat{z}} \left[v(x) - h + \beta V^s(\hat{z}) \right] \\ & s.t.\ h + z = x + (1+\gamma)\hat{z} \end{aligned} \tag{16}$$

where $\hat{z}$ is the seller's real money balance of the next period. In the CM, a seller produces and consumes the CM goods and chooses the amount of money to bring to the next DM. Similar to the buyer's problem, the seller's optimal quantity of the CM consumption x^* satisfies $v'(x^*) = 1$. We also have $W^s(z) = z + W^s(0)$, and the seller's optimal real balance $\hat{z}^*$ does not depend on z.

We then turn to the seller's value function in the DM, which is

$$V^s(\hat{z}) = \max_{p \geq c} \pi(p) + W^s(\hat{z}), \tag{17}$$

where $\pi(p)$ is the seller's profit function, and it does not depend on his money holding. We substitute (17) into (16), and an immediate result for the seller is $\hat{z}^* = 0$ since $1 + \gamma > \beta$. The seller does not bring any money to the DM because he does not want to consume, and the profit is not affected by his real balance.

In the DM, a seller takes the buyer's optimal real balances, the measure of different types of buyers, and the price distribution in the market as given. He chooses a price to maximize the following profit function

$$\pi(p) = (1-\alpha)\left(d^*(p;z_1) - c\frac{d^*(p;z_1)}{p}\right) + 2\alpha\left(1 - F\left(p\right)\right)\left(d^*(p;z_2) - c\frac{d^*(p;z_2)}{p}\right),$$

where z_1 and z_2 represent the buyer's real balance and α is the measure of type-2 buyers in the same period of the seller's price posting problem, and d^* is the buyer's optimal expenditure on the DM goods. With probability $1-\alpha$, the seller is the only one that a buyer visits, and with probability α, he is competing with another seller for the same buyer. In that situation, the seller can have a successful trade only if his price is lower than his competitor's price, which happens with probability $1 - F(p)$. Finally, $d^*(p;z_i) - cd^*(p;z_i)/p$, $i = 1, 2$, denotes the seller's surplus from trade.

We proceed to characterize the upper and lower limit of $F(p)$. Facing the price distribution in the DM, the highest price posted by a seller must be equal to or higher than $\bar{p}$, in which case the seller expects to trade with buyers who only visit him, i.e. type-1 buyers only. Then,

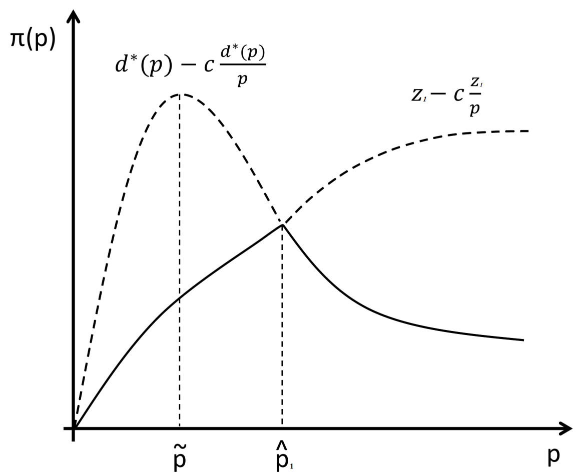

Figure 3: Seller's Profit per Trade in the DM

this seller does not face any competition from other sellers, and his profit function becomes

$$\pi(\bar{p}') = (1-\alpha)\left(d^*(\bar{p}'; z_1) - c\frac{d^*(\bar{p}'; z_1)}{\bar{p}'}\right),$$

where $\bar{p}'$ is the highest price that the seller chooses to post given $F(p)$ and its support $\mathcal{Z}_F$. Notice that this is the same problem faced by every seller, and the optimal choice of $\bar{p}'$ does not depend on distribution F. Therefore, the upper limit of $F(p)$, $\bar{p}$ is determined by the following lemma.

Lemma 3 *Given the optimal money holding of type-1 buyers z_1 and the buyer's optimal expenditure rule $d^*(p; z_1)$, the upper limit of the price distribution $F(p)$ is given by $\bar{p} = \max\{\hat{p}_1, \tilde{p}\}$, where $\tilde{p}$ satisfies $d^*(\tilde{p})u''(d^*(\tilde{p})/\tilde{p})/\tilde{p} + \tilde{p} - c = 0$ and $\hat{p}_1$ is defined in Lemma 1 by $u'(z_1/\hat{p}_1) = \hat{p}_1$.*

Proof. see Appendix. ■

Figure 3 shows the seller's profit as a function of price posted, taking into account the buyer's optimal expenditure in the DM. Notice that the seller's actual profit function is the lower envelope of two separate curves. When the buyer's price elasticity of demand is greater than one, his spending on the DM good decreases with price. Then, without the monetary constraint, the seller chooses to post $\tilde{p}$ to maximize his profit $d^*(p) - cd^*(p)/p$. On the other hand, the buyer actually cannot spend more than the amount of money he carries into the

DM, and the monetary constraint is binding for prices lower than $\hat{p}_1$. For higher prices, the buyer wants to spend less, and the constraint is not binding at all.

When the nominal interest rate is high, it is relatively more costly to hold money for buyers, and it is more likely to have $\tilde{p} < \hat{p}_1$. This case is illustrated in Figure 3. When the interest rate is low and carrying money is less costly, it is possible to have $\tilde{p} > \hat{p}_1$, in which case the seller can earn the highest possible profit at $\tilde{p}$. The next lemma characterizes the price distribution in the DM.

Lemma 4 *Given the buyer's optimal choices on real balance z_1 and z_2 and the measure of type-2 buyers α, the price posting equilibrium distribution $F(p)$ in the DM is uniquely characterized as*

(i) if $\alpha = 0$, $F(p)$ is concentrated at $\bar{p}$.

(ii) if $\alpha = 1$, $F(p)$ is concentrated at c.

(iii) if $\alpha \in (0,1)$, $F(p)$ is nondegenerate and $\mathcal{Z}_F = [\underline{p}, \bar{p}]$ is connected, and for any $p \in \mathcal{Z}_F$,

$$F(p) = 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}; z_1)(\bar{p}-c)\, p}{d^*(p; z_2)(p-c)\,\bar{p}} - \frac{d^*(p; z_1)}{d^*(p; z_2)}\right],$$

where $\bar{p}$ is given in Lemma 3 and $\underline{p}$ satisfies

$$\frac{(1-\alpha)\, d^*(\underline{p}; z_1) + 2\alpha d^*(\underline{p}; z_2)}{(1-\alpha)\, d^*(\bar{p}; z_1)} = \frac{(\bar{p}-c)\,\underline{p}}{\left(\underline{p}-c\right)\bar{p}}. \tag{18}$$

Proof. see Appendix. ■

If every buyer samples just one price, i.e. $\alpha = 0$, each seller behaves like a monopolist, and they all post a price as high as possible in order to extract all the surplus from trade. If every buyer samples two prices, each seller is facing competition from another seller. The seller's price posting game becomes a Bertrand competition, and the competitive price, which is equal to the marginal cost, is posted in equilibrium.

If some of the buyers choose to sample one price while others choose two, a certain degree of competition is introduced among sellers. When a single seller decides which price to post, he faces the trade-off between profit per trade and expected trade volume. If the seller posts a higher price, he gets more profit from one trade, but it is more likely for him to lose the buyer that visits him. If the seller posts a lower price, he has a larger chance to win a buyer but can only get less profit from the trade. This trade-off makes sellers indifferent within an interval of prices and generates a nondegenerate price distribution. If we focus on a symmetric equilibrium in which all sellers choose the same strategy, every seller posts price p with probability $f(p)$, and $f(p) = dF(p)/dp$.

If more buyers search harder by sampling two prices, $F(p)$ increases. The upper limit of the price distribution does not change, since it is determined by type-1 buyer's money holding alone, while the lower limit decreases due to more intense competition. The price dispersion measured as the length of the support of $F(p)$ increases. Increased search behavior intensifies competition among sellers; thus, it is more likely for a buyer to get a relatively low price, and in general the average price level gets lower.

If the nominal interest rate increases and type-1 buyers bring less money to the DM, $\bar{p}$ increases, but the effect on $F(p)$ depends on σ. When σ is less than one and the buyer's demand elasticity is greater than one, sellers respond by increasing the highest price in the market. Then, by the seller's equal profit condition, the overall price level in the DM rises and $F(p)$ decreases.

3.3 Equilibrium

Before defining equilibrium, we close the model by characterizing the buyer's entry decision. When a buyer enters the CM, he first needs to choose to be a type-1 or type-2 buyer, and then makes other optimal choices based on his initial decision on type. The economy can only reach an equilibrium if the marginal buyer is indifferent between being type-1 or type-2. Given an existing composition of the buyer's population, there is not a single buyer who wants to change his type. That is, we need to have $W_1^b(z) = W_2^b(z)$ in order to make $\alpha^* \in (0,1)$. Plug in (10) and (11) and we define

$$\begin{aligned} \Phi(\alpha) &= \int_{\underline{p}}^{\bar{p}} \left[u\left(\frac{d^*(p;\hat{z}_2^*)}{p} \right) - d^*(p;\hat{z}_2^*) \right] \mathrm{d}\left[1-(1-F(p))^2\right] \\ &\quad - \int_{\underline{p}}^{\bar{p}} \left[u\left(\frac{d^*(p;\hat{z}_1^*)}{p} \right) - d^*(p;\hat{z}_1^*) \right] \mathrm{d}F(p) - i\left(\hat{z}_2^* - \hat{z}_1^*\right) \end{aligned} \tag{19}$$

to be the gain of observing two prices instead of one. Then, $W_1^b(z) = W_2^b(z)$ is equivalent to $\Phi(\alpha) = k$.

A buyer can get a better deal by sampling one more price, but he also needs to pay the extra opportunity cost of carrying more money. In order to have an interior solution of α, the above gain needs to be equal to the cost k. Intuitively, if the gain from sampling the second price is less than k, we have $\alpha^* = 0$ and all buyers are type-1. If the gain is larger than the cost, $\alpha^* = 1$ and all buyers choose to be type-2. Notice that even though α does not enter (19) directly, it appears in the expression of the price distribution $F(p)$ in equilibrium.

Definition 1 *A stationary monetary equilibrium (SME) is a profile* $\{F^*, z_1^*, z_2^*, x^*, h^*, d^*, \alpha^*\}$ *satisfying the following conditions:*

1. Given d^**,* z_1^**,* z_2^**, and* α^**, sellers post profit-maximizing prices in the DM and the resulting price distribution* F^* *is determined by Lemma 4;*

2. Given F^**,* d^**, and* α^**,* z_1^**,* z_2^**,* x^**, and* h^* *solve the buyer's problem in the CM;*

3. Given F^**,* z_1^**, and* z_2^**, the buyer's optimal spending rule in the DM* d^* *satisfies Lemma 1;*

4. Given F^**,* d^**,* z_1^**, and* z_2^**, buyers optimize the number of price samplings in the DM and* α^* *is the resulting measure of type-2 buyers.*

In equilibrium, money bears value and circulates because buyers can use it as means of payment in the DM and sellers may use it in exchange for consumption goods in the CM. Equilibrium is symmetric in the sense that all ex ante homogeneous sellers post the same price distribution. In general, two kinds of equilibrium may potentially exist: one with a degenerate price distribution in the DM and one with a nondegenerate price distribution. Proposition 1 shows that the first kind of equilibrium does not exist.

Proposition 1 *If* $1 + \gamma > \beta$*, there exists no SME with* $\alpha^* = 0$ *or* $\alpha^* = 1$*.*

Proof. see Appendix. ■

If all the buyers choose to be type-2 and $\alpha = 1$, the equilibrium price distribution becomes degenerate and concentrated at the marginal cost. Then, any type-2 buyer would want to deviate and switch to type-1, since the marginal gain from observing one more price is zero but he has to pay a non-zero cost. The equilibrium then collapses.

If all the buyers sample just one price and $\alpha = 0$, the equilibrium price distribution again becomes degenerate and concentrated at the highest possible price, i.e. the seller's monopoly price. When posting this price, sellers do not take into account the buyer's opportunity cost of holding money, because it is a sunk cost in the DM. Then, when the buyer chooses the optimal real balance in the CM, he finds that the marginal cost of carrying money to the DM is positive while the marginal gain is zero, since all the surplus from DM trade is exploited by the seller. Therefore, the buyer chooses $z_1^* = 0$ and this is no longer a monetary equilibrium. In the next proposition, we establish the existence of a stationary monetary equilibrium with a nondegenerate price distribution

Proposition 2 *For $1 + \gamma > \beta$, there exists $\bar{k} > 0$, and for $k < \bar{k}$, SME with nondegenerate price distribution exists.*

Proof. see Appendix. ■

This proposition is formally proved in the Appendix, and here we discuss the basic intuition behind the argument. First, we show that the price distribution posted by sellers in the DM is decreasing, in the sense of first-order stochastic dominance, with respect to the real balance that type-1 and type-2 buyers are expected to hold. If buyers of either type carry a larger real balance, the monetary constraint in the buyer's optimal expenditure problem in the DM is relaxed. This implies that buyers now have more money to spend when they meet sellers with relatively low prices, and hence low-price sellers get more profit compared to high-price sellers. In order for the equal profit condition to hold for different prices, sellers must shift down the price distribution. While the average number of type-1 buyers served by each seller stays the same, the expected number of type-2 buyers is reduced at low-price sellers relative to high-price sellers. So sellers remain indifferent between low and high prices.

Then, we show that the amount of real balance carried by buyers in the DM decreases with the price distribution posted by sellers. If buyers are facing a lower price distribution, in the sense of first-order stochastic dominance, the probability of being cash constrained when buying the DM good is higher, since they are more likely to meet a seller with low price. As a result, the marginal benefit of carrying one additional unit of real balance into the DM increases, and then buyers simply want to carry more money.

The intuition above shows that the amount of real balance that buyers want to carry is an increasing function of the amount that they are expected to carry by sellers. Moreover, we can show that the buyer's optimal choice of real balance is bounded above, since no one wants to carry more money when it can no longer help to relax buyer's monetary constraint in the DM. Then, from the fixed point theorem in Tarski (1955), we can prove that for a given measure of type-2 buyers $\alpha \in (0,1)$, there exists optimal real balances for both type-1 and type-2 buyers.

Finally, we show that type-2 buyer's surplus from the DM trade is larger than type-1 buyer's surplus. Intuitively, type-2 buyers carry a larger real balance and face a lower transaction price in the DM, and the benefit outweighs the additional cost they suffer from inflation. After paying a positive search cost, the payoff of both types is the same and no one wants to deviate from existing price sampling decision. Therefore, there exists $\alpha^* \in (0,1)$ such that $\Phi(\alpha^*) - k = 0$, which in turn determines the optimal real balances z_1^* and z_2^* for type-1 and type-2 buyers, the DM price distribution F^* resulting from seller's optimal pricing decision, and all the other endogenous variables. Even though we cannot formally prove the uniqueness of equilibrium, we can make a similar argument as in Wright (2010). When there are more than one α that satisfies $\Phi(\alpha) - k = 0$, the biggest α is the one that yields the highest surplus for type-2 buyers, hence the equilibrium measure of type-2 buyers.

4 Quantitative Analysis

In this section, we calibrate the model to match money demand and price dispersion in U.S. data, and quantitatively measure the aggregate effect of inflation on welfare, as well as the effects through each individual channel. We also check the robustness of the results by considering alternative calibration targets and different time periods.

4.1 Calibration

We assume the following functional forms for preferences and production technology:

$$\begin{aligned} CM &: U(x) = A \log x - h \\ DM &: u(q) = B\frac{q^{1-\sigma}}{1-\sigma} \text{ and } c(q) = cq \end{aligned}$$

where $\sigma > 0$ is the relative risk aversion coefficient of u. The utility function in the CM is standard, following the literature on welfare cost of inflation since Cooley and Hansen (1989). Because both parameters A and B characterize the relative size of the CM economy versus the DM, we normalize $B = 1$. The marginal cost of production in the DM satisfies $c = 1$, so that the cost of labor is the same in both markets.

The time period of the model is set to be one year. We choose this length of time in order to compare the results with those in previous studies, and also because of data availability. We need to calibrate three key parameters, the buyer's search cost k and preference parameters σ and A. The first target we use is money demand, a sample of 101 years, from 1900 to 2000, including the not seasonally adjusted nominal GDP and M1[5] and short-term (6 month) commercial paper rate as the nominal interest rate.[6] In the model, money demand, defined as $L(i) = M/PY$, represents real balance as a function of nominal interest rate. Real balance M/P is proportional to the total real output Y with a factor of proportionality $L(i)$, which depends on the opportunity cost of holding money. Per capita real output in the CM is $x^* = A$, and real output per trade in the DM is

$$(1-\alpha^*)\int_{\underline{p}^*}^{\bar{p}^*} \frac{d^*(p; z_1^*)}{p}\mathrm{d}F^* + \alpha^* \int_{\underline{p}^*}^{\bar{p}^*} \frac{d^*(p; z_2^*)}{p}\mathrm{d}\left[1-(1-F^*)^2\right].$$

[5]Nominal GDP is taken from the Historical Statistics of the United States, Colonial Times to Present (1970) and the GDPA series from the Citibase database. Money supply is M1, as of December of each year, and is not seasonally adjusted. It is from the Historical Statistics of the United States (1970), Friedman and Schwartz (1963), and the FRED II database of the Federal Reserve Bank of St. Louis.

[6]From 1900 to 1997, the short-term commercial paper rate is taken from Friedman and Schwartz (1982), Economic Report of the President (1996), and Economic Report of the President (2003). From 1998 to 2000 it is the short-term 90-day AA credit rate from the Federal Reserve Board.

So the money demand function is

$$L(i)=\frac{(1-\alpha^*)\,z_1^*+\alpha^* z_2^*}{2A+(1-\alpha^*)\int_{\underline{p}^*}^{\bar{p}^*}\frac{d^*(p;z_1^*)}{p}\mathrm{d}F^*+\alpha^*\int_{\underline{p}^*}^{\bar{p}^*}\frac{d^*(p;z_2^*)}{p}\mathrm{d}\left[1-(1-F^*)^2\right]}.$$

Concerning the calibration target for search cost k, there are many empirical studies in which the magnitude of price dispersion is measured by relative price variability (RPV), defined as

$$RPV=\left[\int_{\underline{p}}^{\bar{p}}(R-R_{ave})^2\,\mathrm{d}F(p)\right]^{1/2}$$

where $R=\log(p/\int_{\underline{p}}^{\bar{p}}p\mathrm{d}F)$ and $R_{ave}=\int_{\underline{p}}^{\bar{p}}R\mathrm{d}F$. Debelle and Lamont (1997) report an average RPV of 0.035 at the annual inflation rate of 4.3%,[7] and we use it as the target for k in the baseline calibration. Finally, β is set to match annual real interest rate of 4%.

In order to see how calibration results vary with different target values, we also use another target from Parsley (1996) for price dispersion, and it gives an average RPV of 0.0923 at the annual inflation rate of 5.3%.[8] I also consider different calibration strategies to check the robustness of results. As a very common approach in the literature, search cost is calibrated to match the average markup in the DM, which is 30% at the annual nominal interest rate of 5.46%, as reported in Faig and Jerez (2005).[9]

In order to numerically solve the model, we first fix the measure of type-2 buyers α^* and also z_1^* and z_2^*, and compute the uniquely determined price distribution F^* using Lemma 4. Then, we substitute F^* back into (13) and (15) and solve for z_1^* and z_2^*. Finally, we insert both z_1^*, z_2^* and F^* into (19) and search for $\alpha^*\in(0,1)$ that solves $\Phi(\alpha)=k$.

[7] The data used in Debelle and Lamont (1997) includes two balanced panel of CPI data from BLS, 1954-1986 for 19 cities and 14 kinds of goods and services, 1977-1986 for 24 cities and 18 kinds of goods and services, mostly from the retail sector.

[8] Parsley (1996) uses quarterly survey data from Cost of Living Index published by the American Chamber of Commerce Researchers Association, a panel from 1975 to 1992 with 48 cities and 32 kinds of goods and services, mostly from the retail sector.

[9] This data on markup is from the Annual Retail Trade Survey conducted by the U.S. Census Bureau, which is collected in the format of firm surveys. We are not claiming that the retail market is a real economy counterpart of the decentralized market in the model, but this is the best available data we can get. The average rate of annual inflation for the period covered in the sample is 5.46%.

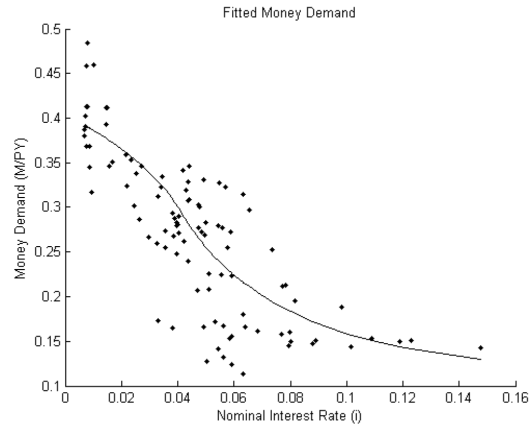

Figure 4: Money Demand: Model and Data

4.2 Results

The baseline calibration gives a very small search cost $k = 0.0043$, and also a very small risk aversion coefficient $\sigma = 0.1181$. The relative size of the CM with respect to the DM implies $A = 0.4916$. The discount factor is $\beta = 0.9615$. The markup in the DM is defined as $\mu_{DM} = \int_{\underline{p}}^{\bar{p}} (p - c) / c \mathrm{d}F(p)$. The baseline calibration predicts the DM markup to be 9.72%, which implies a lot of competition among sellers in the market. The overall markup μ is the average of two markets, weighted by the shares of output, which is 2.4% in the baseline calibration, implying about 25% of total real output being produced in the DM.

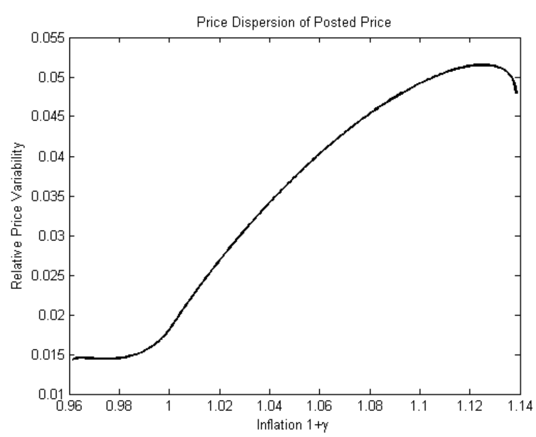

Figure 5: Inflation and Price Dispersion

Figure 4 presents the model-generated money demand function with real data. We use the regression coefficient of inflation on price dispersion as an independent check of the model's fitness. The model implies a coefficient of 0.2784, which falls in the range of 0.115 to 0.393, reported in Debelle and Lamont (1997). So the model does a good job in fitting the empirical data of money demand and price dispersion.

In the baseline calibration, we can only find a unique SME. In equilibrium, as nominal interest rate increases, buyers bring less real balance to the DM. Since $\sigma < 1$, the buyer's price elasticity of demand is greater than one. As they bring less money to the DM, their optimal expenditures decrease and are bound by real balance. The buyer's surplus from trade also decreases, which encourages them to search less. On the other hand, equilibrium price distribution becomes more dispersed with higher inflation, which gives buyers incentive to search harder. Therefore, the model displays a nonmonotonic effect of inflation on search intensity, and hence a nonmonotonic relationship between inflation and price dispersion, as shown in Figure 5.

4.3 Welfare Cost of Inflation

Following Lucas (2000), we measure the welfare cost of increasing annual inflation from zero percent to τ percent by compensating consumption. For any τ, total welfare is given by

$$\begin{aligned}(1-\beta)\,\mathcal{W}(\tau) &= 2\left[v(x^*)-x^*\right]+(1-\alpha^*)\int_{\underline{p}^*}^{\bar{p}^*}\left[u\left(\frac{d^*\left(p;z_1^*\right)}{p}\right)-c\frac{d^*\left(p;z_1^*\right)}{p}\right]\mathrm{d}F^* \\ &\quad +\alpha^*\int_{\underline{p}^*}^{\bar{p}^*}\left[u\left(\frac{d^*\left(p;z_2^*\right)}{p}\right)-c\frac{d^*(p;z_2^*)}{p}\right]\mathrm{d}\left[1-(1-F^*)^2\right]-\alpha^* k, \quad (20)\end{aligned}$$

where $2[v(x^*)-x^*]$ is the total surplus from CM trade. The next two integral terms are the surplus from trade in the DM, subtracting search cost. We can also write the total welfare at zero inflation, with consumption reduced by a factor of Δ in both the CM and DM, as

$$\begin{aligned}(1-\beta)\mathcal{W}_\Delta(0) &= 2\left[v(\Delta x^*)-x^*\right]+(1-\alpha^*)\int_{\underline{p}^*}^{\bar{p}^*}\left[u\left(\frac{\Delta d^*(p;z_1^*)}{p}\right)-c\frac{d^*(p;z_1^*)}{p}\right]\mathrm{d}F^* \\ &\quad +\alpha^*\int_{\underline{p}^*}^{\bar{p}^*}\left[u\left(\frac{\Delta d^*(p;z_2^*)}{p}\right)-c\frac{d^*(p;z_2^*)}{p}\right]\mathrm{d}\left[1-(1-F^*)^2\right]-\alpha^* k.\end{aligned}$$

So we measure the welfare cost of τ percent inflation as the value $1-\Delta_0$ that solves $\mathcal{W}(\tau) = \mathcal{W}_{\Delta_0}(0)$. Every buyer and seller need to give up $1-\Delta_0$ percent of their consumption to be indifferent between the two economies with different inflation rates.

In the baseline calibration, we find the welfare cost of 10% annual inflation is worth 3.23% of consumption in the economy with zero inflation, which is relatively larger compared to earlier findings by Cooley and Hansen (1989) and Lucas (2000), and slightly smaller than those reported by Lagos and Wright (2005), Craig and Rocheteau (2008a), and Rocheteau and Wright (2009). Figure 6 illustrates welfare cost at different levels of inflation, with and without search cost. Obviously, search cost only generates a negligible welfare loss.

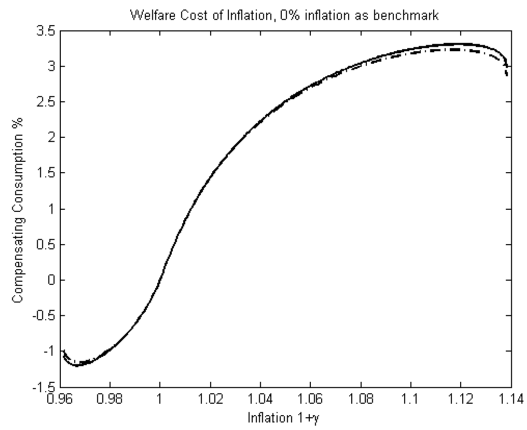

Figure 6: Welfare Cost of Inflation

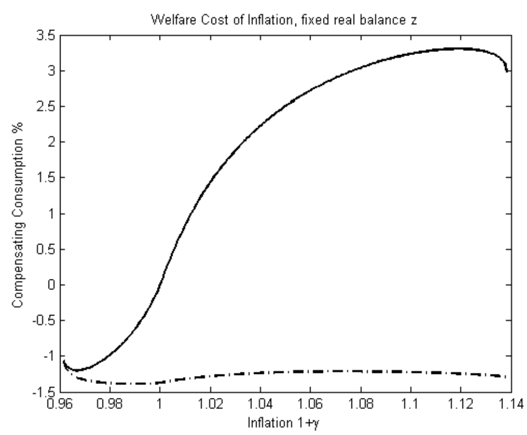

Figure 7: Fixed Real Balance

In this model, inflation affects welfare through three channels: real balance channel, price posting channel, and endogenous search channel. Real balance channel refers to the fact that higher inflation reduces buyer's real balance, and hence consumption in the DM. Because

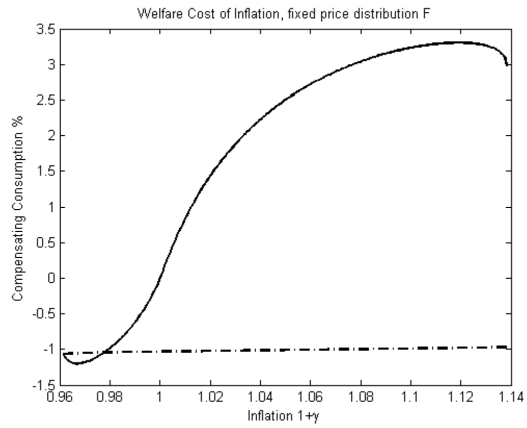

Figure 8: Fixed Price Posting

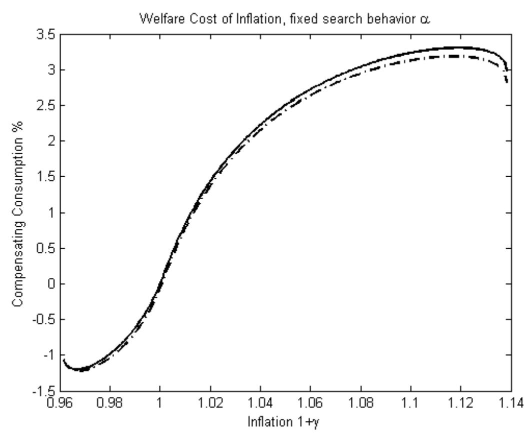

Figure 9: Fixed Search

buyers need to pay a cost to search for prices, sellers have monopolistic power, and can post prices higher than marginal cost. This is price posting channel, which responds to inflation indirectly through other channels. Finally, with higher inflation, buyers search harder for low-price sellers, and this is search channel. Higher search intensity increases competition among sellers, drives down overall price level, and increases the DM consumption and welfare.

The aggregate effect of inflation on welfare is shown in Figure 6. A small deviation from the Friedman rule leads to welfare improvement. Even if it is costless to hold money, sellers still set prices above marginal cost, as Burdett and Judd (1984) show in a cashless economy. If holding money becomes costly, buyers have an incentive to search harder for low prices, and then average transaction price drops and welfare increases. At very low level of inflation, the positive effect on welfare through search channel dominates the negative effect

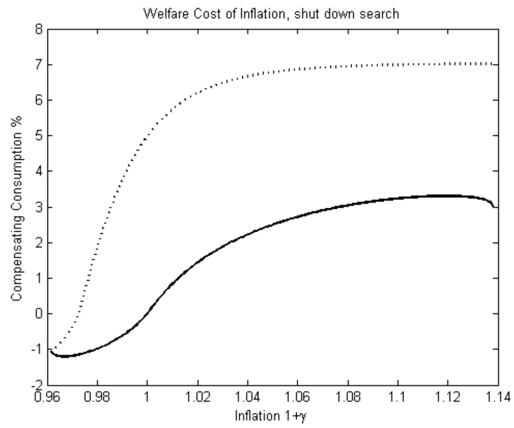

Figure 10: Exogenous Search

through real balance and price posting channel. However, as the cost of holding money keeps increasing, real balance drops quickly, and its first-order effect on consumption soon drives up welfare cost. When almost all buyers sample two prices, the price distribution in the DM is pushed toward marginal cost pricing, and welfare increases marginally. Therefore, the effect of inflation on welfare is nonmonotonic.

In order to study each individual channel and their interactions, we decompose the welfare cost of inflation in the following way. In the baseline calibration, we solve monetary equilibrium at different inflation rates. We keep two channels at their equilibrium values but hold the third channel constant, and recalculate the welfare cost. By comparing with the original values in Figure 6, we can isolate the contribution of the third channel.

First, we hold real balance constant, at the highest possible level, and keep F and α at their equilibrium values. The dashed line in Figure 7 represents the welfare cost of inflation when holding real balance channel constant. The difference between two lines represents the effect of inflation through real balance channel. The welfare cost of 10% annual inflation decreases from 3.23% of zero-inflation consumption to 0.04%. Mitigated by search channel, this value captures the effect mainly through price posting channel, and is similar to what Burstein and Hellwig (2008) find in a New Keynesian menu cost model with monopolistically competitive firms. On the other hand, Figure 8 shows the difference in welfare when holding price distribution F constant, and the welfare cost of 10% inflation drops to 0.15% of consumption. This result represents the effect through real balance and search channel, and is in line with previous findings by Cooley and Hansen (1989) and Lucas (2000), who focus on the opportunity cost of holding money.

The effect through search channel is illustrated in Figure 9, and one observes welfare gain of 0.1% of consumption at 10% inflation rate. This implies that welfare gain through search channel is even less than the welfare loss due to search cost. However, it does not necessarily lead to a trivial effect of endogenous search on welfare. Equation (20) shows that search intensity affects total welfare in two ways. It directly affects total surplus from the DM trade, and indirectly affects real balance and price distribution. Figure 10 presents the change in welfare cost when we completely remove endogenous search decision instead of holding the fraction of type-2 buyers constant, and hence this graph shows the total effect of search channel on welfare. Without endogenous search, the welfare cost of 10% inflation increases to 6.99%. This finding implies that endogenous search lowers the welfare cost by more than 50%.

From the above decomposition exercise, we conclude that the large welfare cost of inflation found in this model is caused by the interaction of real balance and price posting channel. As inflation increases, real balance drops, and the DM consumption decreases. Since the buyer's price elasticity of demand is greater than one, sellers respond by posting even higher prices, which further decreases quantity traded in the DM and welfare. Compared to the holdup problem in generalized Nash bargaining, this interaction of real balance and price posting can generate even larger welfare loss. As pointed out by Lagos and Wright (2005), bargaining can generate a large welfare cost, because buyer's investment in the DM trade, by costly holding real balance, is not fully compensated by the bargaining solution. Unless buyer's bargaining power is one, they choose to underinvest in real balance. In the current model with price posting, when sellers post higher prices in response to lower real balance, it is equivalent to decreasing the buyer's bargaining power in Lagos and Wright (2005), which makes the holdup problem more severe. For example, if we recalibrate the model to match the targets in Lagos and Wright (2005), we find the welfare cost of 10% inflation is about 100% larger than their finding.

Finally, we also consider alternative calibration targets and different data time periods. Table 1 presents calibration results, using price dispersion statistics from Parsley (1996) as

Target 2 and the DM markup of 30% as Target 3, and the welfare cost of 10% inflation. Parsley (1996) finds more dispersed price distribution, and so Target 2 implies a bigger search cost, less competition in the DM, and a larger welfare cost. Similar intuition applies to a larger DM markup.

Table 2 presents a shorter sample of 1959-2000. Generally, calibration results do not change much, and the welfare cost of inflation stays in the same range. Compared to Table 1, both search cost k and the CM preference parameter A are relatively higher. Although it is more costly to search in the DM, agents switch to consume more CM goods, and their aggregate utility from consumption does not necessarily decrease. Because the share of CM trade is much higher than before, the welfare cost of inflation actually becomes smaller.

	Baseline	Target 2	Target 3
k	0.0043	0.0328	0.0057
A	0.4916	1.0134	0.5064
σ	0.1181	0.5326	0.2111
μ_{DM}	9.72%	47.13%	25.64%
μ	2.4%	8.75%	5.52%
$1-\Delta_0$	3.23%	8.31%	7.23%

Table 1. Calibrations with Alternative Targets

	Baseline	Target 2	Target 3
k	0.0089	0.033	0.0061
A	1.1841	1.5488	0.6175
σ	0.1861	0.5262	0.2200
μ_{DM}	8.34%	45.14%	35.32%
μ	1.42%	5.78%	4.79%
$1-\Delta_0$	2.34%	5.08%	6.68%

Table 2. Calibrations for Shorter Period 1959-2000

5 Conclusion

In this paper, we develop a search-theoretic monetary model to incorporate real balance, price posting, and endogenous search. In Head et al (2012), this framework has been proven to fit empirical evidence on price changes well, and is a good tool for welfare analysis. We quantify the welfare cost of inflation, and study different channels through which inflation affects welfare. Calibrating the model to match U.S. data on money demand and price dispersion, we find that the welfare cost of 10% annual inflation is worth 3.23% of consumption in zero inflation economy, which is larger than previous findings in the literature. However, if either real balance or price posting channel is shut down, the welfare cost significantly decreases to less than 0.15% of consumption, and is consistent with previous findings.

The big welfare loss is driven by the interaction of real balance and price posting channel. With higher inflation, buyers respond by holding lower real balance, and sellers respond by posting higher prices. The total effect through two channels decreases consumption and welfare significantly. Buyers search more intensively when inflation gets higher, and endogenous search channel improves welfare mostly by intensifying competition among sellers. The aggregate effect of inflation on welfare through all three channels is nonmonotonic.

This paper models search frictions as the only driving force for buyers to hold money and sellers to post prices, and it is also the source of endogenous price dispersion in equilibrium. This modeling choice is neat and useful in incorporating different channels into one unified environment. There are other frictions that affect seller's price setting behavior and price dispersion, such as menu costs. In a related paper, Burdett and Menzio (2014) show that both menu costs and search frictions contribute to price dispersion, but search frictions are more important than menu costs in matching empirical data. However, it is not clear yet menu costs are less important in the context of studying the welfare cost of inflation, especially with interactions with search frictions, and this is a topic for future research.

References

[1] Aruoba, S. Boragan, and Frank Schorfheide. (2011). "Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-Offs." American Economic Journal: Macroeconomics 3, 60-90.

[2] Bailey, Martin J. (1956). "The Welfare Cost of Inflationary Finance." Journal of Political Economy 64, 93-110.

[3] Benabou, Roland. (1988). "Search, Price Setting and Inflation." Review of Economic Studies 55, 353-376.

[4] Benabou, Roland. (1992). "Inflation and Efficiency in Search Markets." Review of Economic Studies 59, 299-329.

[5] Burdett, Kenneth and Kenneth L. Judd. (1983) "Equilibrium Price Dispersion." Econometrica 51, 955-969.

[6] Burdett, Kenneth and Guido Menzio. (2014). "The (Q,S,s) Pricing Rule." NBER Working Paper 19094.

[7] Burstein, Ariel and Christian Hellwig. (2008). "Welfare Costs of Inflation in a Menu Cost Model." American Economic Review 98, 438-443.

[8] Cooley, F. Thomas and Gary D. Hansen. (1989). "The Inflation Tax in a Real Business Cycle Model." American Economic Review 79, 733-748.

[9] Craig, Ben and Guillaume Rocheteau. (2008a). "Inflation and Welfare: A Search Approach." Journal of Money, Credit and Banking 40, 89-119.

[10] Craig, Ben and Guillaume Rocheteau. (2008b). "State-dependent Pricing, Inflation, and Welfare in Search Economies." European Economic Review 52, 441-468.

[11] Debelle, Guy and Owen Lamont. (1997). "Relative Price Variability and Inflation: Evidence from U.S. Cities." Journal of Political Economy 105, 132-152.

[12] Diamond, Peter A. (1993). “Search, Sticky Prices, and Inflation.” Review of Economic Studies 60, 53-68.

[13] Dutu, Richard, Benoit Julien, and Ian King. (2012). “On the Welfare Gains of Price Dispersion.” Journal of Money, Credit and Banking 44, 757-786.

[14] Faig, Miquel and Belen Jerez. (2005). “A Theory of Commerce.” Journal of Economic Theory 122, 60-99.

[15] Friedman, Milton and Anna Schwartz. (1963). A Monetary History of the United States, 1867-1960. Princeton University Press.

[16] Friedman, Milton and Anna Schwartz. (1982). “The Effect of the Term Structure of Interest Rates on the Demand for Money in the United States.” Journal of Political Economy 90, 201-212.

[17] Head, Allen and Alok Kumar. (2005). “Price Dispersion, Inflation, and Welfare.” International Economic Review 46, 533-572.

[18] Head, Allen, Lucy Qian Liu, Guido Menzio and Randall Wright. (2012). “Sticky Prices: A New Monetarist Approach.” Journal of the European Economic Association 10, 939-973.

[19] Lerner, William. (1970) Historical Statistics of the United States, Colonial Times to 1970. Bureau of the Census, United States Department of Commerce.

[20] Lagos, Ricardo and Randall Wright. (2005). “A Unified Framework for Monetary Theory and Policy Analysis.” Journal of Political Economy 113, 463-484.

[21] Liu, Lucy Qian, Liang Wang and Randall Wright. (2014). “Sticky Prices and Costly Credit.” Mimeo.

[22] Lucas, Robert E., Jr. (2000). “Inflation and Welfare.” Econometrica 68, 247-274.

[23] Parsley, David C. (1996). “Inflation and Relative Price Variability in the Short and Long Run: New Evidence from the United States.” Journal of Money, Credit and Banking 28, 323-341.

[24] Rocheteau, Guillaume and Randall Wright. (2005). “Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium.” Econometrica 73, 175-202.

[25] Rocheteau, Guillaume and Randall Wright. (2009). “Inflation and Welfare in Models with Trading Frictions.” in Monetary Policy in Low Inflation Economies, edited by David E. Altig and Ed Nosal, pp. 89-116. New York: Cambridge University Press.

[26] Shi, Shouyong. (1997). “A Divisible Search Model of Fiat Money.” Econometrica 65, 75-102.

[27] Tarski, Alfred. (1955). “A Lattice Theoretical Fixpoint Theorem and Its Applications.” Pacific Journal of Mathematics 5, 285-309.

[28] Wright, Randall. (2010). “A Uniqueness Proof for Monetary Steady State.” Journal of Economic Theory 145, 382-391.

[29] Wallace, Neil. (2011). “The Mechanism-Design Approach to Monetary Theory.” in Handbook of Monetary Economics 3A, edited by Benjamin M. Friedman and Michael Woodford, pp. 3-23. San Diego: North-Holland.

Appendix

Proof of Lemma 1.

The first order condition of the buyer's unconstrained optimization problem is $u'(d^*/p)/p - 1 = 0$. Notice that $\sigma < 1$ implies $-qu''(q)/u'(q) < 1$, so $u'(d^*/p)/p$ is a decreasing function in p and $\partial d^*(p)/\partial p < 0$. Hence, there exists $\hat{p}_j$ such that $d^*(\hat{p}_j) = z_j$ for $z_j > 0$, $j = 1, 2$. For $p < \hat{p}_j$, $d^*(p) > z_j$ and $u'(z_j/p)/p - 1 > 0$. Buyers want to spend more, but are subject to the constraint $d^*(p; z_j) \leq z_j$. Hence, $d^*(p; z_j) = z_j$ for $p < \hat{p}_j$. For $p \geq \hat{p}_j$, $d^*(p) \leq z_j$ and $d^*(p; z_j) = d^*(p)$. It is straightforward to verify that $\partial \hat{p}_j/\partial z_j < 0$. ■

Proof of Lemma 2.

In this proof, I use $z_j^* > 0$, $j = 1, 2$ to denote the optimal real balance of type-1 and type-2 buyers, respectively. In the following, we focus on type-1 buyers and the proof for type-2 buyers can be derived similarly. We proceed by contradiction and assume that $\hat{p}_1 \leq \underline{p}$ at z_1^*. We first consider the situation in which $\hat{p}_1 < \underline{p}$. If $\hat{p}_1 < \underline{p}$, $d^*(p; z_1) = d^*(p)$ for all $p \in [\underline{p}, \bar{p}]$. We plug it into (12) and omit all the terms unrelated to z_1, and the buyer's optimization problem can be rewritten as

$$L = \max_{z_1} \left\{ -iz_1 + \int_{\underline{p}}^{\bar{p}} \left[u\left(\frac{d^*(p)}{p} \right) - d^*(p) \right] \mathrm{d}F(p) \right\}. \tag{21}$$

The first order condition with respect to z_1 evaluated at z_1^* is $\partial L/\partial z_1^* = -i < 0$, which is a contradiction to z_1^* being the optimal real balance. Then, we consider the situation of $\hat{p}_1 = \underline{p}$. Recall that $\hat{p}_1$ is determined by z_1 through $u'(z_1/\hat{p}_1) = \hat{p}_1$, so z_1^* satisfies $\hat{p}(z_1^*) = \underline{p}$. We want to solve for $\partial L/\partial z_1^*$. When z_1 approaches z_1^* from below, $\hat{p}(z_1)$ approaches $\underline{p}$ from above since $\partial \hat{p}_1/\partial z_1 < 0$, and we can rewrite (21) as

$$\begin{aligned} L &= \max_{z_1} \left\{ -iz_1 + \int_{\hat{p}_1}^{\bar{p}} \left[u\left(\frac{d^*(p)}{p} \right) - d^*(p) \right] \mathrm{d}F(p) \right. \\ &\quad \left. + \int_{\underline{p}}^{\hat{p}_1} \left[u\left(\frac{z_1}{p} \right) - z_1 \right] \mathrm{d}F(p) \right\}, \end{aligned}$$

thus $\lim_{z_1 \to z_{1-}^*} \partial L/\partial z_1 = \lim_{z_1 \to z_{1-}^*} \{-i + \int_{\underline{p}}^{\hat{p}_1} [u'(z_1/p)/p - 1] \mathrm{d}F(p)\} = -i < 0$. On the other hand, when z_1 approaches z_1^* from above, $\hat{p}(z_1)$ approaches $\underline{p}$ from below, and $\lim_{z_1 \to z_{1+}^*} \partial L/\partial z_1 =$

$-i < 0$. However, the fact that z_1^* is the optimal real balance implies $\lim_{z_1 \to z_{1-}^*} \partial L / \partial z_1 \geq 0$ and $\lim_{z_1 \to z_{1+}^*} \partial L / \partial z_1 \leq 0$. This is a contradiction. Therefore, we must have $\hat{p}_1 > \underline{p}$ for type-1 buyers. ■

Proof of Lemma 3.

First, notice that when $p = \bar{p}$ is posted, only type-1 buyers purchase from this seller, since all the other prices are below $\bar{p}$. Given the expression for type-1 worker's cutoff price $\hat{p}_1$ in Lemma 1, there are two cases: $\bar{p} \leq \hat{p}_1$ or $\bar{p} > \hat{p}_1$. If $\bar{p} \leq \hat{p}_1$, $d^*(\bar{p}; z_1) = z_1$, and $\pi(\bar{p}) = z_1 (1 - \alpha^*) (1 - c/\bar{p})$. A seller wants to choose a price as high as possible in the feasible range, and he posts $\bar{p} = \hat{p}_1$. In the other case with $\bar{p} > \hat{p}_1$, $d^*(\bar{p}; z_1) = d^*(\bar{p})$ satisfying $u'(d^*(\bar{p})/\bar{p}) = \bar{p}$. The seller wants to choose $\bar{p}$ such that

$$(1 - \frac{c}{\bar{p}}) \frac{\partial d^*(\bar{p})}{\partial \bar{p}} + d^*(\bar{p}) \frac{c}{\bar{p}^2} = 0, \tag{22}$$

which is the first order condition of the seller's profit maximization problem. We can derive $\partial d^*(\bar{p}) / \partial \bar{p}$ from $u'(d^*(\bar{p})/\bar{p}) = \bar{p}$, and substitute it into (22). Hence, if $\bar{p} > \hat{p}_1$, $\bar{p} = \tilde{p}$ where $\tilde{p}$ is given by

$$\frac{d^*(\tilde{p})}{\tilde{p}} u''(\frac{d^*(\tilde{p})}{\tilde{p}}) + \tilde{p} - c = 0.$$

Therefore, seller wants to post the upper limit $\bar{p} = \max\{\hat{p}_1, \tilde{p}\}$. ■

Proof of Lemma 4.

If $\alpha = 0$, the seller's profit function is

$$\pi(p) = d^*(p; z_1) - c \frac{d^*(p; z_1)}{p}. \tag{23}$$

According to Lemma 3, there is a unique price that maximizes $\pi(p)$, so every seller posts $\bar{p}$. If a seller deviates by posting $\bar{p}' = \bar{p} + \varepsilon$, where $\varepsilon > 0$, his profit per trade decreases since $\bar{p}$ maximizes (23) and his trade volume stays the same. Similarly, if a seller deviates to $\bar{p} - \varepsilon$, his profit drops without an increase in trade volume. Therefore, there is no incentive for any seller to deviate away from $\bar{p}$.

If $\alpha = 1$, it is clearly an equilibrium that every seller posts $p = c$. There is no incentive to post a price lower than c, since that yields a negative profit. On the other hand, if a seller deviates and posts $c+\varepsilon$, his profit $\pi(c+\varepsilon)$ is equal to zero since $F(c+\varepsilon)=1$ and he loses all the buyers. Next, I want to argue that this is the only equilibrium of the seller's price posting game. If there is another $F(p)$ concentrated at $p'>c$, a seller has incentive to lower the posted price by ε, i.e. he wants to post $p'-\varepsilon$. In this way, he can trade with a buyer for sure even though his profit from the trade decreases a little. A discrete increase in the trading probability makes up for the infinitesimal drop of the profit, and the seller's expected profit increases. Hence, there is a profitable deviation and another degenerate $F(p)$ does not exist. If there is another nondegenerate $F(p)$, its support $\mathcal{Z}_F$ is connected. This conclusion follows directly from Lemma 1 in Burdett and Judd (1983). $\pi(p)$ must be the same for all $p\in\mathcal{Z}_F$, and in particular $\pi(p)=\pi(\bar{p})=[2-2F(\bar{p})][d^*(\bar{p};z_2)-cd^*(\bar{p};z_2)/\bar{p}]=0$ since $F(\bar{p})=1$. However, for any p such that $F(p)\in(0,1)$, $\pi(p)=[2-2F(p)][d^*(p;z_2)-cd^*(p;z_2)/p]>0$. This is a contradiction. Therefore, $F(p)$ concentrated at c is the unique equilibrium price distribution in the seller's price posting game.

If $\alpha^*\in(0,1)$, any $F(p)$ concentrated at $p\in[\underline{p},\bar{p}]$ cannot be a price posting equilibrium distribution. On the one hand, seller can always increase the price, hence increase the profit, while still keeping those buyers who only sample his price. On the other hand, seller can also lower his price infinitesimally, and get a jump in the trading probability. Hence, $F(p)$ must be nondegenerate if $\alpha^*\in(0,1)$. Again from Lemma 1 in Burdett and Judd (1983), we know $F(p)$ is continuous with connected support. For any $p\in[\underline{p},\bar{p}]$, we must have $\pi(p)=\pi(\bar{p})$, which implies

$$
\begin{aligned}
& (1-\alpha)\left[d^*(p;z_1)-c\frac{d^*(p;z_1)}{p}\right]+2\alpha\left[1-F(p)\right]\left[d^*(p;z_2)-c\frac{d^*(p;z_2)}{p}\right] \\
= \; & (1-\alpha)\left[d^*(\bar{p};z_1)-c\frac{d^*(\bar{p};z_1)}{\bar{p}}\right].
\end{aligned}
$$

The above equation determines a unique $F(p)$ for each p. In particular, $\pi(\underline{p})=\pi(\bar{p})$ determines $\underline{p}$, which satisfies

$$
(1-\alpha)\,d^*(\bar{p};z_1)\left(1-\frac{c}{\bar{p}}\right)=\left[(1-\alpha)\,d^*\left(\underline{p};z_1\right)+2\alpha d^*\left(\underline{p};z_2\right)\right]\left(1-\frac{c}{\underline{p}}\right).
$$

■

Proof of Proposition 1.

First, we consider the case of $\alpha^* = 0$. Suppose an SME exists for an economy with $1+\gamma > \beta$. Because $\alpha^* = 0$, the equilibrium price distribution in the DM, $F^*(p)$ must be concentrated at $\bar{p}$.

When $\sigma < 1$, $\bar{p} = \max\{\hat{p}, \tilde{p}\}$. If $\tilde{p} \geq \hat{p}$, $d^*(\bar{p}; z_1^*) = d^*(\tilde{p})$. The type-1 buyer's optimal real balance z_1^* maximizes

$$L = -(1+\gamma-\beta)z_1 + \beta\left[u\left(\frac{d^*(\tilde{p})}{\tilde{p}}\right) - d^*(\tilde{p})\right].$$

Immediately, we have $\partial L/\partial z_1^* = -(1+\gamma-\beta) < 0$, and $z_1^* = 0$. This contradicts the existence of a monetary equilibrium. If $\hat{p} > \tilde{p}$, $d^*(\bar{p}; z_1^*) = z_1^*$, and z_1^* maximizes

$$L = -(1+\gamma-\beta)z_1 + \beta\left[u\left(\frac{z_1}{\hat{p}}\right) - z_1\right].$$

Then, $\partial L/\partial z_1^* = -(1+\gamma-\beta) + \beta[u'(z_1^*/\hat{p})/\hat{p} - 1] = -(1+\gamma-\beta) < 0$, since $u'(z_1^*/\hat{p}) = \hat{p}$ by Lemma 1. Hence, $z_1^* = 0$ and it is again a contradiction. Therefore, there does not exist an SME with $\alpha^* = 1$. ■

Proof of Proposition 2.

We want to show that there exists $\alpha^* \in (0,1)$, $z_1^* > 0$, $z_2^* > 0$, and F^* satisfying $\Phi(\alpha^*) = k$, (13), (15), and Lemma (4). We first show that given $\alpha \in (0,1)$, there exists z_1^*, z_2^*, and F^* such that type-1 buyers choose $z_1 = z_1^*$ and type-2 buyers choose $z_2 = z_2^*$ given the price distribution $F^*(p; z_1^*, z_2^*)$. Then, we show that we can find $\alpha^* \in (0,1)$ such that $\Phi(\alpha^*) = k$. We proceed in four steps.

Claim 1 *Let $\underline{z}^*$ and $\bar{z}^*$ be defined by*

$$\hat{p}(\underline{z}^*) = \tilde{p} \text{ and } \hat{p}(\bar{z}^*) = \underline{p},$$

where $\tilde{p}$ is defined in Lemma (3) and $\underline{p}$ is defined in

$$\frac{(1+\alpha)\, d^*(\underline{p})}{(1-\alpha)\, d^*(\tilde{p})} = \frac{(\tilde{p}-c)\, \underline{p}}{(\underline{p}-c)\, \tilde{p}}. \tag{24}$$

Define $Z^* = (z_1^*, z_2^*)$. Then, for $Z_0^* = (z_{10}^*, z_{20}^*)$ and $Z_1^* = (z_{11}^*, z_{21}^*)$ such that $0 < Z_0^* <$

$Z_1^* \leq (\bar{z}^*, \bar{z}^*)$, $F(p; Z_0^*)$ first-order stochastically dominates $F(p; Z_1^*)$. For Z_0^* and Z_1^* such that $(\bar{z}^*, \bar{z}^*) \leq Z_0^* < Z_1^*$, then $F(p; Z_0^*) = F(p; Z_1^*)$.

Proof. First, consider the case in which $(\bar{z}^*, \bar{z}^*) \leq Z_0^* < Z_1^*$. Then, the price distribution is

$$F(p; Z^*) = 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}^*)(\bar{p}^* - c)\,p}{d^*(p)(p-c)\,\bar{p}^*} - 1\right]$$

with support $\left[\underline{p}^*, \bar{p}^*\right]$, where $\bar{p}^* = \tilde{p}$ and $\underline{p}^* = \underline{p}$ as defined in (24). In this case, $F(p; Z_0^*) = F(p; Z_1^*)$ for all Z_0^* and Z_1^* such that $(\bar{z}^*, \bar{z}^*) \leq Z_0^* < Z_1^*$.

Second, consider the case $\underline{z}^* \leq z_{1i}^* \leq \bar{z}^* \leq z_{2i}^*$ for $i = 0, 1$ and $Z_0^* < Z_1^*$. In this case, the price distribution is

$$F(p; Z^*) = \begin{cases} 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}^*)(\bar{p}^*-c)p}{d^*(p)(p-c)\bar{p}^*} - 1\right] & \text{if } p \in \left[\hat{p}(z_1^*), \bar{p}^*\right] \\ 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}^*)(\bar{p}^*-c)p}{d^*(p)(p-c)\bar{p}^*} - \frac{z_1^*}{d^*(p)}\right] & \text{if } p \in \left[\underline{p}^*, \hat{p}(z_1^*)\right] \end{cases}$$

with support $\left[\underline{p}^*, \bar{p}^*\right]$, where $\bar{p}^* = \tilde{p}$ and $\underline{p}^*$ is the solution to

$$\frac{(1-\alpha)\,z_1^* + 2\alpha d^*\left(\underline{p}^*\right)}{(1-\alpha)\,d^*(\bar{p}^*)} = \frac{(\bar{p}^* - c)\,\underline{p}^*}{\left(\underline{p}^* - c\right)\bar{p}^*}.$$

Since $Z_0^* < Z_1^*$, we have $\bar{p}^*(Z_0^*) = \bar{p}^*(Z_1^*) = \bar{p}^*$, $\underline{p}^*(Z_1^*) < \underline{p}^*(Z_0^*)$, $\hat{p}(z_{11}^*) < \hat{p}(z_{10}^*)$, and $\hat{p}(z_{21}^*) < \hat{p}(z_{20}^*)$. It is also straightforward to verify that given α and F, for $Z_i^* = (z_{1i}^*, z_{2i}^*)$, $z_{1i}^* < z_{2i}^*$, for $i = 0, 1$. For $p \geq \hat{p}(z_{10}^*)$, $F(p; Z_0^*) = F(p; Z_1^*)$. For $p \in [\hat{p}(z_{11}^*), \hat{p}(z_{10}^*)]$, then $F(p; Z_0^*) < F(p; Z_1^*)$ since $z_{10}^* < d^*(p)$. For $p \in \left[\underline{p}^*(Z_0^*), \hat{p}(z_{11}^*)\right]$, we also have $F(p; Z_0^*) < F(p; Z_1^*)$ because $z_{10}^* < z_{11}^*$. As for $p \in \left[\underline{p}^*(Z_1^*), \underline{p}^*(Z_0^*)\right]$, we still have $F(p; Z_0^*) = 0 < F(p; Z_1^*)$. For $p \leq \underline{p}^*(Z_1^*)$, $F(p; Z_0^*) = F(p; Z_1^*) = 0$. Hence, $F(p; Z_0^*)$ first-order stochastically dominates $F(p; Z_1^*)$.

Third, consider the case $(\underline{z}^*, \underline{z}^*) \leq Z_0^* < Z_1^* \leq (\bar{z}^*, \bar{z}^*)$. The price distribution is

$$F(p; Z^*) = \begin{cases} 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}^*)(\bar{p}^*-c)p}{d^*(p)(p-c)\bar{p}^*} - 1\right] & \text{if } p \in \left[\hat{p}(z_1^*), \bar{p}^*\right] \\ 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}^*)(\bar{p}^*-c)p}{d^*(p)(p-c)\bar{p}^*} - \frac{z_1^*}{d^*(p)}\right] & \text{if } p \in \left[\hat{p}(z_2^*), \hat{p}(z_1^*)\right] \\ 1 - \frac{1-\alpha}{2\alpha}\left[\frac{d^*(\bar{p}^*)(\bar{p}^*-c)p}{z_2^*(p-c)\bar{p}^*} - \frac{z_1^*}{z_2^*}\right] & \text{if } p \in \left[\underline{p}^*, \hat{p}(z_2^*)\right] \end{cases}$$

with support $\left[\underline{p}^*, \bar{p}^*\right]$, where $\bar{p}^* = \tilde{p}$ and $\underline{p}^*$ is the solution to

$$\frac{(1-\alpha)\, z_1^* + 2\alpha z_2^*}{(1-\alpha)\, d^*(\bar{p}^*)} = \frac{(\bar{p}^* - c)\, \underline{p}^*}{\left(\underline{p}^* - c\right) \bar{p}^*}.$$

Since $Z_0^* < Z_1^*$, we have $\bar{p}^*(Z_0^*) = \bar{p}^*(Z_1^*) = \bar{p}^*$, $\underline{p}^*(Z_1^*) < \underline{p}^*(Z_0^*)$, $\hat{p}(z_{11}^*) < \hat{p}(z_{10}^*)$, and $\hat{p}(z_{21}^*) < \hat{p}(z_{20}^*)$. There are two possibilities regarding the relationship between $\hat{p}(z_{11}^*)$ and $\hat{p}(z_{20}^*)$, and then we need to discuss two different situations, $\hat{p}(z_{21}^*) < \hat{p}(z_{20}^*) < \hat{p}(z_{11}^*) < \hat{p}(z_{10}^*)$ and $\hat{p}(z_{21}^*) < \hat{p}(z_{11}^*) < \hat{p}(z_{20}^*) < \hat{p}(z_{10}^*)$. We discuss the first situation in the following and the second situation can be verified similarly. For $p \geq \hat{p}(z_{10}^*)$, $F(p; Z_0^*) = F(p; Z_1^*)$. For $p \in [\hat{p}(z_{11}^*), \hat{p}(z_{10}^*)]$, then $F(p; Z_0^*) < F(p; Z_1^*)$ since $z_{10}^* < d^*(p)$. For $p \in [\hat{p}(z_{20}^*), \hat{p}(z_{11}^*)]$, $F(p; Z_0^*) < F(p; Z_1^*)$ because $z_{10}^* < z_{11}^*$. For $p \in [\hat{p}(z_{21}^*), \hat{p}(z_{20}^*)]$, $F(p; Z_0^*) < F(p; Z_1^*)$ since $z_{10}^*/z_{20}^* < z_{11}^*/d^*(p)$. For $p \in [\underline{p}^*(Z_0^*), \hat{p}(z_{21}^*)]$, because $z_{10}^* < z_{11}^*$ and $z_{20}^* < z_{21}^*$, $F(p; Z_0^*) < F(p; Z_1^*)$. As for $p \in [\underline{p}^*(Z_1^*), \underline{p}^*(Z_0^*)]$, $F(p; Z_0^*) = 0 < F(p; Z_1^*)$. Hence, $F(p; Z_0^*)$ first-order stochastically dominates $F(p; Z_1^*)$ in this case.

Finally, consider the case $z_{1i}^* \leq \underline{z}^* \leq z_{2i}^* \leq \bar{z}^*$ for $i = 0, 1$ and $Z_0^* < Z_1^*$, which is the same as the case $z_{1i}^* \leq z_{2i}^* \leq \underline{z}^* \leq \bar{z}^*$. The price distribution is

$$F(p; Z^*) = \begin{cases} 1 - \frac{1-\alpha}{2\alpha}\left[\frac{z_1^*(\bar{p}^*-c)p}{d^*(p)(p-c)\bar{p}^*} - \frac{z_1^*}{d^*(p)}\right] & \text{if } p \in [\hat{p}(z_2^*), \bar{p}^*] \\ 1 - \frac{1-\alpha}{2\alpha}\left[\frac{z_1^*(\bar{p}^*-c)p}{z_2^*(p-c)\bar{p}^*} - \frac{z_1^*}{z_2^*}\right] & \text{if } p \in \left[\underline{p}^*, \hat{p}(z_2^*)\right] \end{cases}$$

with support $\left[\underline{p}^*, \bar{p}^*\right]$, where $\bar{p}^* = \hat{p}(z_1^*)$ and $\underline{p}^*$ is the solution to

$$1 + \frac{2\alpha z_2^*}{(1-\alpha)\, z_1^*} = \frac{(\bar{p}^* - c)\, \underline{p}^*}{\left(\underline{p}^* - c\right) \bar{p}^*}.$$

Since $Z_0^* < Z_1^*$, we have $\bar{p}^*(Z_1^*) = \hat{p}(z_{11}^*) < \hat{p}(z_{10}^*) = \bar{p}^*(Z_0^*)$, $\underline{p}^*(Z_1^*) < \underline{p}^*(Z_0^*)$, and $\hat{p}(z_{21}^*) < \hat{p}(z_{20}^*)$. For $p \in [\bar{p}^*(Z_1^*), \bar{p}^*(Z_0^*)]$, $F(p; Z_0^*) < F(p; Z_1^*) = 1$. For $p \in [\hat{p}(z_{20}^*), \bar{p}^*(Z_1^*)]$, we can verify that $F(p; Z_0^*) < F(p; Z_1^*)$ because $z_{10}^* < z_{11}^*$. For $p \in [\hat{p}(z_{21}^*) < \hat{p}(z_{20}^*)]$, $F(p; Z_0^*) < F(p; Z_1^*)$ since $z_{10}^*/z_{20}^* < z_{11}^*/d^*(p)$. For $p \in [\underline{p}^*(Z_0^*), \hat{p}(z_{21}^*)]$, $F(p; Z_0^*) < F(p; Z_1^*)$ since $z_{10}^* < z_{11}^*$ and $z_{20}^* < z_{21}^*$. For $p \in [\underline{p}^*(Z_1^*), \underline{p}^*(Z_0^*)]$, $F(p; Z_0^*) = 0 < F(p; Z_1^*)$. Therefore, $F(p; Z_0^*)$ first-order stochastically dominates $F(p; Z_1^*)$ again in this case and we have proved Claim 1.

Claim 2 *Given α and $F(p; z_1^*, z_2^*)$, let the unique solution for z_1 and z_2 in household's CM problem be $z_i = \varphi_i(z_1^*, z_2^*)$ for $i = 1, 2$. Then, φ_i have the following properties:*
(i) $\forall\ z_{i0}^*,\ z_{i1}^*$ such that $0 < z_{i0}^* < z_{i1}^* \le \bar{z}^*$, $\varphi_i(z_{10}^*, z_{20}^*) \le \varphi_i(z_{11}^*, z_{21}^*)$, for $i = 1, 2$.
(ii) $\forall\ z_{i0}^*,\ z_{i1}^*$ such that $\bar{z}^* \le z_{i0}^* < z_{i1}^*$, $\varphi_i(z_{10}^*, z_{20}^*) \le \varphi_i(z_{11}^*, z_{21}^*)$, for $i = 1, 2$.
(iii) $\forall\ z_i^* > 0$, $\varphi_i(z_1^*, z_2^*) \in [\underline{\varphi}_i, \bar{\varphi})$, where $\bar{\varphi} = \bar{z}^*$ and $\underline{\varphi}_i > 0$, for $i = 1, 2$.

Proof. Given the measure of type-2 buyers α and the price distribution $F(p; z_1^*, z_2^*)$, the equilibrium conditions for real balances z_1 and z_2 are

$$\int_{\underline{p}(z_1^*, z_2^*)}^{\hat{p}(z_1)} \left[u'\left(\frac{z_1}{p}\right)\frac{1}{p} - 1\right] \mathrm{d}F\left(p; z_1^*, z_2^*\right) = i \tag{25}$$

and

$$\int_{\underline{p}(z_1^*, z_2^*)}^{\hat{p}(z_2)} \left[u'\left(\frac{z_2}{p}\right)\frac{1}{p} - 1\right] \mathrm{d}\left[1 - \left(1 - F\left(p; z_1^*, z_2^*\right)\right)^2\right] = i. \tag{26}$$

Let $\chi_1\left(z_1; z_1^*, z_2^*\right)$ denote the left-hand side of (25) and $\chi_2(z_2; z_1^*, z_2^*)$ denote the left-hand side of (26). One can easily show that

$$\begin{aligned}
\lim_{z_1 \to 0} \chi_1\left(z_1; z_1^*, z_2^*\right) &= \infty \\
\lim_{z_2 \to 0} \chi_2\left(z_2; z_1^*, z_2^*\right) &= \infty
\end{aligned}$$

and $\chi_i\left(z_i; z_1^*, z_2^*\right)$ strictly decreases in z_i, $\forall z_i \in (0, \hat{z}_i)$ where $\hat{p}(\hat{z}_i) = \underline{p}\left(z_1^*, z_2^*\right)$, for $i = 1, 2$. Also notice that $\chi_i\left(z_i; z_1^*, z_2^*\right) = 0$, $\forall z_i \ge \hat{z}_i$. Hence, there is a unique solution $z_1 = \varphi_1(z_1^*, z_2^*)$ to (25) and a unique solution $z_2 = \varphi_2(z_1^*, z_2^*)$ to (26) with $\varphi_i(z_1^*, z_2^*) \in (0, \hat{z}_i)$, for $i = 1, 2$.

Consider $0 < Z_0^* < Z_1^* \le (\bar{z}^*, \bar{z}^*)$. According to Claim 1, $F(p; Z_0^*)$ first-order stochastically dominates $F(p; Z_1^*)$, and then $1 - (1 - F(p; Z_0^*))^2$ first-order stochastically dominates $1 - (1 - F(p; Z_1^*))^2$. Since $u'\left(z_i/p\right)/p - 1$ decreases in z_i for $i = 1, 2$, we have $\chi_1\left(z_1; Z_0^*\right) \le \chi_1\left(z_1; Z_1^*\right)$ and $\chi_2\left(z_2; Z_0^*\right) \le \chi_2\left(z_2; Z_0^*\right)$, which implies $\varphi_1(Z_0^*) < \varphi_1(Z_1^*)$ and $\varphi_2(Z_0^*) < \varphi_2(Z_1^*)$. One can also easily verify that $\varphi_i(Z_0^*) \ge \underline{\varphi}_i$ for some $\underline{\varphi}_i > 0$, $i = 1, 2$.

Finally, consider the case $(\bar{z}^*, \bar{z}^*) \le Z_0^* < Z_1^*$. According to Claim 1, $F(p; Z_0^*) = F(p; Z_1^*)$ and $1 - (1 - F(p; Z_0^*))^2 = 1 - (1 - F(p; Z_1^*))^2$ Consequently, we have $\chi_i\left(z_i; Z_0^*\right) \le \chi_i\left(z_i; Z_1^*\right)$, and hence $\varphi_i(Z_0^*) = \varphi_i(Z_1^*) < \bar{z}^*$, for $i = 1, 2$.

Claim 3 *Given $\alpha \in (0, 1)$, there exists $z_i^* \in [\underline{\varphi}_i, \bar{\varphi})$ such that $\varphi_i(z_1^*, z_2^*) = z_i^*$, for $i = 1, 2$.*

Proof. Define $\varphi(z_1^*, z_2^*) = (z_1, z_2) = (\varphi_1(z_1^*, z_2^*), \varphi_2(z_1^*, z_2^*))$. Claim (2) then implies that $\varphi(z_1^*, z_2^*)$ is increasing in both arguments and $\varphi(z_1^*, z_2^*) \in [\underline{\varphi}_1, \bar{\varphi}) \times [\underline{\varphi}_2, \bar{\varphi})$, $\forall (z_1^*, z_2^*) \in [\underline{\varphi}_1, \bar{\varphi}] \times [\underline{\varphi}_2, \bar{\varphi}]$. By Tarski's fixed point theorem, there exists $(z_1^*, z_2^*) \in [\underline{\varphi}_1, \bar{\varphi}) \times [\underline{\varphi}_2, \bar{\varphi})$ such that $(z_1^*, z_2^*) = \varphi(z_1^*, z_2^*)$, i.e. $z_1^* = \varphi_1(z_1^*, z_2^*)$ and $z_2^* = \varphi_2(z_1^*, z_2^*)$. This proves existence. Claim (2) also implies that $\forall z_i^* \leq \underline{\varphi}_i$, $\varphi_i(z_1^*, z_2^*) \geq \underline{\varphi}_i$, and $\forall z_i^* \geq \bar{\varphi}$, $\varphi_i(z_1^*, z_2^*) < \bar{\varphi} = \bar{z}^*$, for $i = 1, 2$. Therefore, if z_1^* and z_2^* satisfy $z_1^* = \varphi_1(z_1^*, z_2^*)$ and $z_2^* = \varphi_2(z_1^*, z_2^*)$, we must have $z_1^* \in [\underline{\varphi}_1, \bar{\varphi})$ and $z_2^* \in [\underline{\varphi}_2, \bar{\varphi})$.

Claim 4 *$\exists \bar{k} > 0$ such that for $k \in (0, \bar{k})$, there exists $\alpha^* \in (0,1)$ such that $\Phi(\alpha^*) = k$.*

Proof. Claim (3) implies that given $\alpha \in (0,1)$, we can find $z_1^* < z_2^*$, hence $\hat{p}(z_2^*) < \hat{p}(z_1^*)$, and $F^* = F(p; z_1^*, z_2^*)$ that solve the household's problem and the firm's problem. Then, $\Phi(\alpha)$ can be defined as

$$\begin{aligned}\Phi(\alpha) &= \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_2^*)}{p}\right) - d^*(p; z_2^*)\right] \mathrm{d}\left[1 - (1 - F^*)^2\right] \\ &\quad - \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_1^*)}{p}\right) - d^*(p; z_1^*)\right] \mathrm{d}F^* - i(z_2^* - z_1^*).\end{aligned}$$

Notice that F^* first-order stochastically dominates $1 - (1 - F^*)^2$ and $u(d^*(p; z_1^*)/p) - d^*(p; z_1^*)$ is decreasing in p, and then

$$\int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_1^*)}{p}\right) - d^*(p; z_1^*)\right] \mathrm{d}\left[1 - (1 - F^*)^2\right] > \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_1^*)}{p}\right) - d^*(p; z_1^*)\right] \mathrm{d}F^*.$$

This implies that

$$\begin{aligned}\Phi(\alpha) &> \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_2^*)}{p}\right) - d^*(p; z_2^*)\right] \mathrm{d}\left[1 - (1 - F^*)^2\right] \\ &\quad - \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_1^*)}{p}\right) - d^*(p; z_1^*)\right] \mathrm{d}\left[1 - (1 - F^*)^2\right] - i(z_2^* - z_1^*) \\ &= \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_2^*)}{p}\right) - d^*(p; z_2^*) - u\left(\frac{d^*(p; z_1^*)}{p}\right) + d^*(p; z_1^*)\right] \mathrm{d}\left[1 - (1 - F^*)^2\right] \\ &\quad - i(z_2^* - z_1^*)\end{aligned}$$

Then, we plug in (26) and get

$$
\begin{aligned}
\Phi(\alpha) &> \int_{\underline{p}^*}^{\bar{p}^*} \left[u\left(\frac{d^*(p; z_2^*)}{p}\right) - d^*(p; z_2^*) - u\left(\frac{d^*(p; z_1^*)}{p}\right) + d^*(p; z_1^*) \right] \mathrm{d}\left[1 - (1 - F^*)^2\right] \\
&\quad - \int_{\underline{p}^*}^{\hat{p}(z_2^*)} \left[u'\left(\frac{z_2^*}{p}\right)\frac{1}{p} - 1 \right] (z_2^* - z_1^*)\, \mathrm{d}\left[1 - (1 - F^*)^2\right] \\
&> \int_{\underline{p}^*}^{\hat{p}(z_2^*)} \left[u\left(\frac{z_2^*}{p}\right) - z_2^* - u\left(\frac{z_1^*}{p}\right) + z_1^* \right] \mathrm{d}\left[1 - (1 - F^*)^2\right] \\
&\quad - \int_{\underline{p}^*}^{\hat{p}(z_2^*)} \left[u'\left(\frac{z_2^*}{p}\right)\frac{1}{p} - 1 \right] (z_2^* - z_1^*)\, \mathrm{d}\left[1 - (1 - F^*)^2\right].
\end{aligned}
$$

Since $u(z/p) - z$ is concave in z, for $z_1^* < z_2^*$, we have

$$
u\left(\frac{z_2^*}{p}\right) - z_2^* - \left[u\left(\frac{z_1^*}{p}\right) - z_1^* \right] \geq (z_2^* - z_1^*) \left[u'\left(\frac{z_2^*}{p}\right)\frac{1}{p} - 1 \right].
$$

The above equation then implies $\Phi(\alpha) > 0$ for $\alpha \in (0, 1)$.

Since $\Phi(\alpha)$ is a continuous function defined on $[0, 1]$, it must attain a maximum value, denoted as $\bar{k}$. On the other hand, it is straightforward to verify that $\Phi(0) = 0$ and $\Phi(1) = 0$, then $\exists \tilde{\alpha} \in (0, 1)$ such that $\Phi(\tilde{\alpha}) = \bar{k}$. Therefore, for $k \in (0, \bar{k})$, there exists $\alpha^* \in (0, 1)$ such that $\Phi(\alpha^*) - k = 0$. ■