Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Dealers' Insurance, Market Structure, And Liquidity

Francesca Carapella and Cyril Monnet

2017-119

Please cite this paper as:
Carapella, Francesca, and Cyril Monnet (2017). "Dealers' Insurance, Market Structure, And Liquidity," Finance and Economics Discussion Series 2017-119. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.119.

Dealers' Insurance, Market Structure, And Liquidity*

Francesca Carapella†

Federal Reserve Board of Governors

Cyril Monnet‡

University of Bern & Study Center Gerzensee

October 31, 2017

Abstract

We develop a parsimonious model to study the equilibrium structure of financial markets and its efficiency properties. We find that regulations aimed at improving market outcomes can cause inefficiencies. The welfare benefit of such regulation stems from endogenously improving market access for some participants, thus boosting competition and lowering prices to the ultimate consumers. Higher competition, however, erodes profits from market activities. This has two effects: it disproportionately hurts more efficient market participants, who earn larger profits, and it reduces the incentives of all market participants to invest ex-ante in efficient technologies. The general equilibrium effect can therefore result in a welfare cost to society. Additionally, this economic mechanism can explain the resistance by some market participants to the introduction of specific regulation which could appear to be unambiguously beneficial.

Keywords: Liquidity, dealers, insurance, central counterparties

JEL classification: G11, G23, G28

*We thank Ted Temzelides, Alberto Trejos, Pierre-Olivier Weill, Randall Wright, as well as audiences at the second African Search & Matching Workshop, the 2012 Money, Banking, and Liquidity Summer Workshop at the Chicago Fed, the 2012 Banking and Liquidity Wisconsin Conference, and the 2013 meetings of the Society for Economic Dynamics for very helpful comments.

†francesca.Carapella@frb.gov

‡cyril.monnet@vwi.unibe.ch

1 Introduction

Many markets operate through market makers or similar intermediaries. Two elements are most important for market making, counterparty risk and the cost of holding inventories. Both elements have been or will be affected by the G20-led reform to the over-the-counter (OTC) derivatives market following the financial crisis. As part of this reform, G20 Leaders agreed in 2009 to mandate central clearing of all standardized OTC derivatives. Currently, although central clearing rates have increased globally, there still is a significant proportion of OTC derivatives that is not cleared centrally.[1] As the regulatory framework is being implemented, and as changes in the infrastructure landscape for trading and settlement take place (e.g. due to Brexit), little is known about the effects of these reforms on the structure of the markets in which they are implemented.

In this paper, we analyze the effects of introducing measures aimed at reducing counterparty risk and improving liquidity, such as central clearing (FSB [2017], pg. 7), on the structure of financial markets. One may expect that initiatives aimed at reducing such risk would bring uncontested benefits. However, in line with the theory of the second best, we show that such initiatives may to some extent "back-fire": market makers may take actions that can yield to inefficient outcomes. For instance, they may have too little incentive to innovate. Our results are consistent with empirical findings on the effects of mandatory central clearing for Credit Default Swaps indexes in the United States. Studying separately the effects of each implementation phase of the Dodd Frank reform, Loon and Zhong [2016] find that the effect of central clearing on a measure of transaction-level spread is significantly different according to the category of market participants affected by the reform. In particular, central clearing is correlated with an increase in spreads for swap dealers and with a decrease in spreads for commodity pools and all other swap market participants.[2] In our model, the final general equilibrium effect of introducing an insurance mechanism against counterparty risk (e.g. central clearing) crucially depends on features of the market participants involved.

We use a simple set-up with market makers intermediating trades between buyers and sellers. Dealers are heterogeneous, as they can be more or less efficient at making mar-

[1] See FSB [2017] *Review of OTC derivatives market reforms*, June 2017, pg. 2-14, Figures 2,3.

[2] See Loon and Zhong [2016], Table 10 and Appendix A.2.1, pg. 667-9.

kets. For a (fixed) cost they can invest into a market making technology which lowers their expected cost of intermediating a transaction between buyers and sellers. This technology stands in for more efficient balance sheet management, a larger network of investors, etc.

Once they whether or not to invest, dealers post and commit to bid and ask prices. Buyers and sellers sample dealers randomly and decide whether to trade at the posted bid or ask, or whether they should carry on searching for a dealer next period. The search friction implies that the equilibrium bid-ask spreads will be positive. Also, even less efficient dealers will be active because buyers and sellers may be better off accepting an offer which they know is not the best on the market rather than waiting for a better offer. Therefore, our search friction defines the structure of the market measured by how many and which dealers are operating, and its liquidity measured by the distribution of bid-ask spreads.

Contrary to Duffie et al. [2005], dealers are exposed to the risk of having to hold inventories. To make markets, dealers have to accommodate buy-orders with sell-orders. However we assume that buyers (and sellers) can default after placing their orders. If dealers can perfectly forecast how many buyers will default, they will just acquire fewer assets. Otherwise they may find themselves with too many assets in inventory for longer than expected. For simplicity, we make the extreme assumption that market makers cannot sell the asset if the buyer defaults. In this sense, the asset is bespoke. Dealers maximize their expected profit by posting bid-ask spreads that depend on the inventory risk as well as on their cost of intermediating transactions (in the model, a dealer's idiosyncratic transaction cost). In particular less efficient dealers may find optimal to stay out of market making activities.

We then analyze the effects of a set of regulations aimed at lowering counterparty risk and improving pricing[3] on the liquidity and the structure of intermediated markets. In particular, we focus on (i) the measure of active dealers, buyers and sellers, (ii) the share of the market that each dealer services, and (iii) the equilibrium distribution of bid-ask spreads. Such a comprehensive characterization of the equilibrium allows the identification of gainers and losers from such regulations.

[3] See the Commodity Futures Trading Commission (CFTC) reports on the swap regulation introduced by Title VII, Part II of the Dodd-Frank Wall Street Reform and Consumer Protection Act, and FSB [2017] pg.3, 22.

1.1 Model and results

We model regulations as a reduction in the severity of counterparty risk which affects dealers' inventory risk. Intuitively, dealers should benefit from a reduction or elimination of inventory risk. This could be implemented by the introduction of central clearing in the market for an asset, for example, a more liquid secondary market for the asset, a better functioning of the inter-dealer market as in Duffie et al. [2005], or the use of an insurance mechanism between market makers (e.g. credit default swaps (CDS) market).[4]

Everything else constant, a reduction in counterparty risk will result in a reduction of the bid-ask spread. Two distinct mechanisms are responsible for the lower spread, a direct and an indirect one. First, facing a lower default risk, dealers prefer to charge a lower mark up per transaction and execute a larger volume. Second, lower counterparty risk induces less efficient dealers to enter the market thus increasing competition. As a result, more buyers and sellers are served, and the measure of dealers active on the market increases. More efficient dealers however have a lower profit because they lose some market share to lesser efficient dealers. In fact, the most efficient dealers would prefer some counterparty risk as long as other dealers are not fully insured against such risk.

We also analyze the impact of a reduction in counterparty risk on dealers' incentive to adopt a market-making technology that lowers their ex-ante intermediation cost. Protection against risk can induce dealers to opt for a worse market making technology, which can be inefficient. As discussed, reducing risk allows less efficient dealers to enter the market. This additional competition reduces profits of more efficient dealers (ex-post). Since the benefit of becoming a more efficient dealer is smaller, the incentives to invest in the better market-making technology decrease. If the fixed cost of the better technology is too high, dealers will prefer not to invest to become ex-ante more efficient. In turn the entire pool of dealers become worse. This adversely impacts buyers and sellers who face worse terms of trade on average. As a consequence, the introduction of a seemingly beneficial insurance mechanism against counterparty risk reduces welfare of buyers and sellers, unless dealers receive a transfer that compensate their investment into more efficient market making technologies.

This paper thus makes two contributions: first it offers a perspective that can explain

[4] In Appendix D we provide a full characterization of the mapping from a reduction in counterparty risk to the introduction of central clearing.

the opposition of some dealers to tighter regulation, such as mandatory central clearing for all standardized derivatives traded OTC.[5] Second, it argues that forcing the adoption of seemingly beneficial regulation has consequences for the incentives of some market participants. This effect, in general equilibrium, can ultimately have adverse effects on other agents' welfare.

1.2 Related literature

The literature on the microstructure of markets is large and has been mostly interested with explaining bid-ask spreads. It is not our intention to cover this literature here, and we refer the interested reader to O'hara [1995]. Among the first to study the inventory problem of market makers are Amihud and Mendelson [1980]. Here, we are not interested in the inventory management problem per-se as much as in how the cost of managing inventories affects liquidity. In particular, we normalize the optimal size of inventory to zero and we analyze how the probability to experience deviations from this optimal inventory level affects liquidity.

Our paper, by focusing on the effect of competition on the adoption of better market-making technologies, is also related to Dennert [1993] and Santos and Scheinkman [2001]. Following the seminal contribution of Kyle [1985], Dennert [1993] analyzes the effect of competition on bid-ask spreads and liquidity, and shows that liquidity traders might prefer to trade with a monopolist market maker. Santos and Scheinkman [2001] study the effects of competing platforms when there is a risk of default. They show that a monopolist intermediary may ask for relatively little guarantee against the risk of default.

The papers that are most related to ours are the equilibrium search models of Spulber [1996] and Rust and Hall [2003], which we extend by introducing inventory risk through the default of buyers. Duffie et al. [2005] present an environment where market makers are able to trade their inventory imbalances with each other after each trading rounds. Therefore, market makers never carry any inventory in equilibrium. We depart from Duffie et al. [2005] by assuming that market markets may have to hold inventories and we study the effect of

[5] Dodd-Frank Act for example, Dudley [March 22, 2012]. European financial markets legislation has also been moving in the same direction.

regulations, whose goal is to make market makers closer to the set-up in Duffie et al. [2005], on the structure of the market. In an environment similar to Duffie et al. [2005], Weill [2007] shows that competitive market makers offer the socially optimal amount of liquidity, provided they have access to sufficient capital to hold inventories. Weill [2011] shows that if market makers face a capacity constraint on the number of trades which they can conduct, then delays in reallocating assets among investors emerge, thus creating a time-varying bid ask spread, widening and narrowing as market makers build up and unwind their inventories. In contrast to the last papers, we analyze the incentives of dealers to enter market making activities in the first place. In this respect, our paper is also related to Atkeson et al. [2015], who study the incentives of ex-ante heterogenous banks to enter and exit an OTC market. This allows Atkeson et al. [2015] to identify the banks which behave as end users versus the banks which intermediate transactions, and thus behave as dealers. In contrast, we analyze the impact of current OTC market reforms on dealers' entry and investment decisions, and on the efficiency of the resulting equilibrium allocation.

Section 2 describes the basic structure of the model. To understand the basic mechanism underlying our main results, we analyze the equilibrium with no counterparty risk (i.e. settlement fails) in Section 3 and the equilibrium with counterparty risk/settlement fails in Section 4. Section 5 contains our result about the incentives of market makers to invest in a more efficient market making technology ex-ante. Section 6 concludes.

2 A Model of Dealers and Risk

We base our analysis on a modified version of the equilibrium search models in Spulber [1996] and Rust and Hall [2003]. The presentation of the model follows closely the one in Rust and Hall [2003]. There are three types of agents: traders, who can be either buyers or sellers, and dealers. To be consistent with Spulber [1996] and Rust and Hall [2003], we will also sometimes refer to buyers as consumers and to sellers as producers. Buyers and sellers cannot trade directly an asset and all trades must be intermediated by dealers.

There is a continuum $[0, 1]$ of heterogeneous, infinitely-lived, and risk neutral buyers, sellers, and dealers.[6] A seller of type v can sell at most one unit of the asset at an opportunity cost v. A buyer of type v can hold at most one unit of the asset and is willing to pay at most v to hold it.

Dealers face no counterparty risk in Rust and Hall [2003], as dealers' clients exit the market after they settle their claim. Contrary to Rust and Hall [2003], we introduce counterparty risk for dealers by assuming that buyers first place *orders* with dealers, but then exit the market with probability λ, *before* they have the chance to settle their orders. A buyer who exits the market is replaced with a new buyer whose v is drawn from the uniform distribution over $[0, 1]$. We do not consider strategic default and λ is exogenous. This is akin to the risk that a counterparty goes bust for reasons that are independent of its trading activities, and we refer to it as settlement risk. Contrary to buyers, sellers always settle their orders.[7]

In and of itself, this type of counterparty risk is aggregate and not interesting: There is nothing a dealer can do to insure against it. So we also assume that dealers face idiosyncratic risk: Nature does not allocate buyers perfectly across dealers who can be in two states, $s = 1$ and $s = -1$. In state $s = 1$, a dealer has a measure $\lambda - \varepsilon$ of his buyers exiting the market, while in state $s = -1$ a measure $\lambda + \varepsilon$ of his buyers exit. This default shock is independent of whether the buyers placed an order at the bid-ask spread posted by the dealer. Dealers cannot observe state s before it occurs: They only observe the actual measure of buyers exiting the market once that is realized. This shock is i.i.d. and each state occurs with probability 1/2 , so that there is no aggregate uncertainty. Notice also that on average buyers exit the market before settlement with probability λ.

At time $t = 0$, the initial distribution of types of buyers and sellers is $v \sim U[0, 1]$. Since the type of newborn agents is drawn randomly over the same distribution, then the distribution of types will also be $U[0, 1]$ in all subsequent periods $t = 1, 2, 3, ...$ Therefore $U[0, 1]$ is the unique invariant distribution of types in each subsequent period $t = 1, 2, 3, ...$

There is a continuum of dealers indexed by their trading cost k which is the marginal cost of taking a seller's order *before* the seller actually pays for the good.[8] Trading costs

[6] In Appendix E we analyze a version of this model with risk averse traders.

[7] This asymmetry between buyers and seller is not substantial. Analogous results would arise if sellers exited the market before settlement.

[8] This introduces an asymmetry regarding the cost of dealing with a buyer or a seller, which can be

are uniformly distributed over the interval $[\underline{k}, 1]$, where $\underline{k}$ is the marginal cost of the most efficient dealer.

In equilibrium, only dealers who can make a profit will operate a trading post and there will be a threshold level of trading cost, $\bar{k} \leq 1$, such that no dealer with a cost greater than $\bar{k}$ operates a post. A dealer of type $k \in \left[\underline{k}, \bar{k}\right]$ chooses a pair of bid-ask prices $(b(k), a(k))$ that maximizes his expected discounted profits. A dealer is willing to buy the asset at price $b(k)$ from a seller and is willing to sell the asset at the ask price $a(k)$. We consider a stationary equilibrium so that $b(k)$ and $a(k)$ will be constant through time.

Buyers and sellers engage in search for a dealer. Each period, if he decides to search, a trader gets a price quote from a random dealer. Since dealers post stationary bid and ask prices depending on their types, traders face distributions $F(a)$ and $G(b)$ of ask and bid prices. These distributions are equilibrium objects. Traders discount the future at rate β.

Timing, also shown in Figure 1, is as follows: At time 0, dealers $k \in \left[\underline{k}, \bar{k}\right]$ choose a bid and ask quote. $\forall t \geq 0$, buyers and sellers decide whether they want to search or not. If so, they contact a dealer at random, and they either accept the quoted price or keep searching. If they agree, they place an order to buy/sell a unit of the asset. Then each buyer exits with probability λ. Moreover, if a dealer is in state $s \in \{-1, 1\}$, then a measure $\lambda - s\varepsilon$ of his buyers exit before settlement. Finally, settlement occurs: Each operating dealer receives assets from the sellers who placed an order and delivers one asset to each of the $(1 - \lambda + s\varepsilon)$ buyers who settle their orders. Dealers must dispose of the surplus of assets.[9]

The main difference from Spulber [1996] is that buyers do not give up on future options by trading in a given period. In Spulber [1996], buyers exit the market after they trade. Here, trading today does not exclude traders from future trading opportunities. Hence, their trading decision is simpler in Spulber [1996], and dealers do not *compete* but behave as monopolists. A common feature between Spulber [1996] and our set-up is that each active dealer has a higher probability of intermediating funds whenever few dealers operate. This

justified in real contracts as the cost of handling the good underlying the contract. The result would not be substantially modified if we introduced a handling cost of the buyer as well, k^b as long as $k^b < k$. Here we set $k^b = 0$. For financial contracts, this is the cost of designing the contracts.

[9]We could assume that dealers gets some value $\bar{p}$ for each unit of asset they hold and we normalize $\bar{p} = 0$, so that the asset fully depreciates in the hand of the dealers. This low holding-value also stands in for high regulatory costs of holding some assets (such as higher capital requirements).

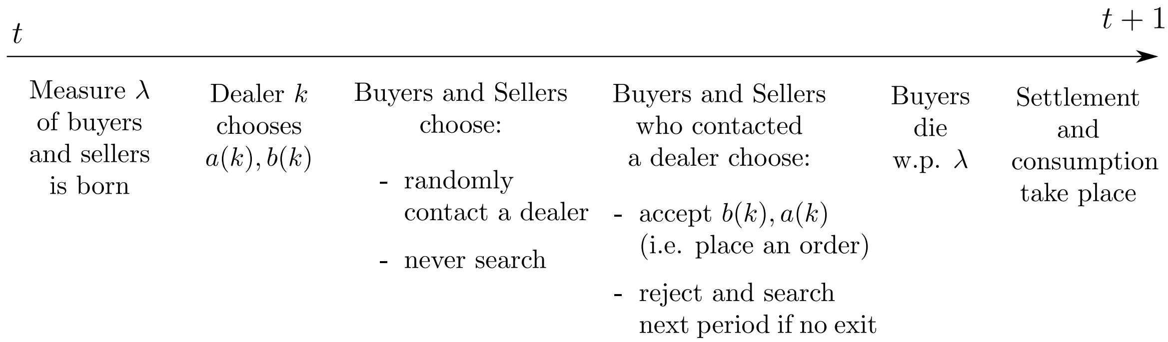

Figure 1: Timing

is key to our result.

3 No settlement risk

To gain some intuition, in this section we study the benchmark economy where there is no settlement risk so that $\lambda = 0$. The decision of buyers/consumers is simply to accept the selected ask price a whenever $v \geq a$ and reject otherwise. Their payoff is

$$V_c(v) = \int_{\underline{a}}^{v} (v - a) dF(a) + \beta V_c(v)$$

where $\underline{a}$ is the lowest ask price. The decision of sellers/producers is to accept the selected bid price b whenever $v \leq b$ and reject otherwise. Their payoff is

$$V_p(v) = \int_{v}^{\bar{b}} (b - v) dG(b) + \beta V_p(v)$$

Dealers that post an ask-price a face the following demand

$$D(a) = \frac{1}{N} \int_{a}^{1} dv = \frac{1}{N}(1 - a) \tag{1}$$

where N is the measure of active dealers. Only those consumers with a value greater than the posted price will accept the offer. Similarly, dealers that post a bid-price b face the following demand

$$S(b) = \frac{1}{N} \int_0^b dv = \frac{1}{N} b \tag{2}$$

A dealer of type k maximizes his profit by choosing a and b, subject to the resource constraint, or

$$\Pi(k) = \max_{a,b} \{aD(a) - (b+k)S(b)\}$$

subject to $D(a) \leq S(b)$. The resource constraint will bind, so that $b = 1 - a$ and a dealer chooses a to maximize

$$\Pi(k) = (1-a)(2a-1-k)$$

with solution

$$a(k) = \frac{3+k}{4} \tag{3}$$

$$b(k) = \frac{1-k}{4} \tag{4}$$

Notice that, as in the models of Spulber [1996] and Rust and Hall [2003], the distribution of bid and ask prices are uniform on $\left[a(0), a(\bar{k})\right]$ and $\left[b(\bar{k}), b(0)\right]$ because the bid and ask prices are linear and the distribution of dealer cost is uniform.

In equilibrium, all dealers with intermediation cost k such that $\Pi(k) \geq 0$ will be active. Therefore, all dealers with $k \leq \bar{k}$, where $\bar{k}$ is defined so that $\Pi(\bar{k}) = 0$, will be active. So the measure of active dealers is $N = \bar{k}$. It is easy to see that $\bar{k} = 1$ and that $a(\bar{k}) = 1$ and $b(\bar{k}) = 0$. Therefore the least efficient dealer is indifferent between operating and staying out of the market. In fact, dealer $\bar{k}$ would face a measure zero demand at the price $a(\bar{k}) = 1$. Any dealer $k < \bar{k} = 1$ makes strictly positive profits:

$$\Pi(k) = \frac{(1-k)^2}{8N} = \frac{(1-k)^2}{8\bar{k}}.$$

Then we can find the extremes of the support of the bid and ask price distributions:

$$\overline{a} = a(\bar{k}) = \frac{3+\bar{k}}{4} = 1 \qquad \underline{a} = a(0) = \frac{3}{4}$$

$$\bar{b} = b(0) = \frac{1}{4} \qquad\qquad \underline{b} = b(\bar{k}) = \frac{1-\bar{k}}{4} = 0$$

Clearly, each dealer charges its monopoly price, as there is no competition: The bid/ask prices posted by *other* dealers do not influence the decision of traders to accept or reject the price they obtain as traders can anyway search again next period, independently of their decision today. So, contrary to the model in Spulber [1996], agents do not forfeit the option of getting a better deal tomorrow if they accept the proposed deal today. Since dealers charge the monopoly price, even inefficient dealers can make profits, which implies that they have the incentive to enter the market: Hence we should expect that the equilibrium number of active dealers is too high relative to what a planner would choose. We analyze this next.

To define the optimal number of dealers, we now define the surplus of dealers, consumers and producers as a function of $\bar{k}$. Total economy-wide profits, or surplus of dealers, are:

$$\begin{aligned} S_d(\bar{k}) = \int_0^{\bar{k}} \Pi(k)dk &= \int_0^{\bar{k}} \frac{(1-k)^2}{8\bar{k}}dk \\ &= \frac{3-(3-\bar{k})\bar{k}}{24} \end{aligned}$$

which are always decreasing in $\bar{k} \leq 1$. The surplus of consumers is:

$$\begin{aligned} S_c(\bar{k}) &= \int_{a(0)}^{1} \left[\int_{a(0)}^{a(\bar{k})\vee v} \frac{(v-a)}{a(\bar{k})-a(0)} da \right] dv \\ &= \frac{(3-(3-\bar{k})\bar{k})}{96} = \frac{S_d(\bar{k})}{4} \end{aligned}$$

Hence, $S_c(\bar{k})$ is always decreasing in $\bar{k}$. Finally, the surplus of producers is

$$\begin{aligned} S_p(\bar{k}) &= \int_{0}^{b(0)} \left[\int_{b(\bar{k})\wedge v}^{b(0)} \frac{(b-v)}{b(0)-b(\bar{k})} db \right] dv \\ &= \frac{(3-(3-\bar{k})\bar{k})}{96} = \frac{S_d(\bar{k})}{4} \end{aligned}$$

Hence $S_p(\bar{k})$ is always decreasing in $\bar{k}$. Therefore, as expected, neither dealers, nor consumers or producers benefit from the entry of relatively inefficient dealers. Given that intermediation is needed, the best solution is to have only the most efficient dealers, those with $k = 0$, intermediate all trades. Notice that this is the case because the most efficient dealer charges the same bid and ask prices independent of the presence of other dealers. This is not true in a model like Spulber [1996], where even the most efficient dealers may wish to lower their price when other dealers are operating. In the next section we introduce settlement risk.

4 Settlement risk

In this section we introduce settlement risk for dealers. A settlement fail occurs when the consumer fails to collect and pay for his buyer order. We assume that this happens on average with probability λ, so that, on average, a measure λ of consumers will fail to settle. However, dealers are also subject to an idiosyncratic settlement shock s with support $S = \{-1, +1\}$ and probability density $\Pr[s = -1] = \Pr[s = +1] = \frac{1}{2}$. This settlement shock describes our notion of counterparty risk: given $\varepsilon \in (0, \lambda)$, a dealer experiences a fraction $\lambda + \varepsilon$ of its consumers failing to settle in state $s = -1$ and a fraction $\lambda - \varepsilon$ failing to settle in state $s = 1$.[10] The cost of settlement fails for dealers is that they still have to honor their obligations toward sellers. The cost of settlement fails for buyers is that they cannot consume the good. We assume that the settlement shock is i.i.d across dealers and across time. We interpret an increase (decrease) in dealers' idiosyncratic settlement risk as an increase (decrease) in ε.

The decision problems of consumers and producers are the same as in the previous section, so that $D(a) = \frac{(1-a)}{N}$ and $S(b) = \frac{b}{N}$. Dealers' decision problem is:

$$\begin{aligned}
\Pi(k; \lambda, \varepsilon) &= \max_{\{a,b\}} E_s \left\{ a \left(1 - \lambda + s\varepsilon\right) D\left(a\right) - \left(b + k\right) S\left(b\right) \right\} & (5) \\
& \quad s.t. \left(1 - \lambda + s\varepsilon\right) D\left(a\right) \leq S\left(b\right) \quad \forall s \in \{-1, 1\} & (6)
\end{aligned}$$

The resource constraint (6) binds when $s = 1$. Therefore

$$S(b) = \left(1 - \lambda + \varepsilon\right) D\left(a\right) \equiv \lambda_\varepsilon D(a).$$

[10] We can extend this to a symmetrically distributed ε around $[-\bar{\varepsilon}, \bar{\varepsilon}]$, where $\bar{\varepsilon} < \lambda$ and $E(\varepsilon) = 0$. Then everything below holds with $\varepsilon = \bar{\varepsilon}$.

Notice that dealers expect to have to deliver $(1-\lambda)D(a)$ assets. However, dealers have to purchase more securities than they expect will be necessary, as they have to satisfy their buy orders in all possible states. Hence, settlement risk implies that dealers over-buy the asset. Substituting out for $D(a)$ and $S(b)$ yields:

$$\Pi(k;\lambda,\varepsilon) \quad = \quad \max_{\{a\}} \left\{a\,(1-\lambda) - \left[\lambda_\varepsilon\,(1-a)+k\right]\lambda_\varepsilon\right\} \frac{1}{N}\,(1-a) \tag{7}$$

Taking the number of operating dealers as given, Figure 2 shows the profit of a dealer when $\varepsilon = 0$ and as ε increases: the direct effect of increasing risk is to reduce dealers' profits. Thus, dealers' best response is to increase their ask price (i.e. $a'(\varepsilon) > 0$). The mechanism driving this result is intuitive: If he posts ask price a, a dealer receives $D(a)$ buy orders but expects only $(1-\lambda)D(a)$ buyers to collect the asset and pay for it. However, he needs to buy sufficient assets to cover effective demand in state $s = 1$. Because such demand increases in ε, an increase in ε reduces dealers' profits. To account for this, dealers adjust their ask price upwards. As a consequence they face fewer buy orders, which, in turn, results in lower effective demand in state $s = 1$.

The first order conditions to dealers' decision problem imply:

$$a(k) \quad = \quad 1 - \frac{1-\lambda-k\lambda_\varepsilon}{2\,(1-\lambda+\lambda_\varepsilon^2)} = \frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)} \tag{8}$$

$$b(k) \quad = \quad \lambda_\varepsilon(1-a(k)) = \lambda_\varepsilon \frac{1-\lambda-k\lambda_\varepsilon}{2\,(1-\lambda+\lambda_\varepsilon^2)} \tag{9}$$

It is worth emphasizing the effect of increasing risk on the bid-ask spread. Since the ask price is increasing with risk, dealers do not need to serve as many consumers as before, so they should decrease their bid price to purchase a lower quantity of the asset. However, notice the factor λ_ε which multiplies $1-a(k)$ in (9): the indirect effect of higher settlement risk is that dealers have to over-buy the security, which pushes the bid price up. The overall effect on the bid price is therefore uncertain, and depends on which effects dominates. It turns out that if λ and ε are sufficiently small, then the bid price will increase in the risk of

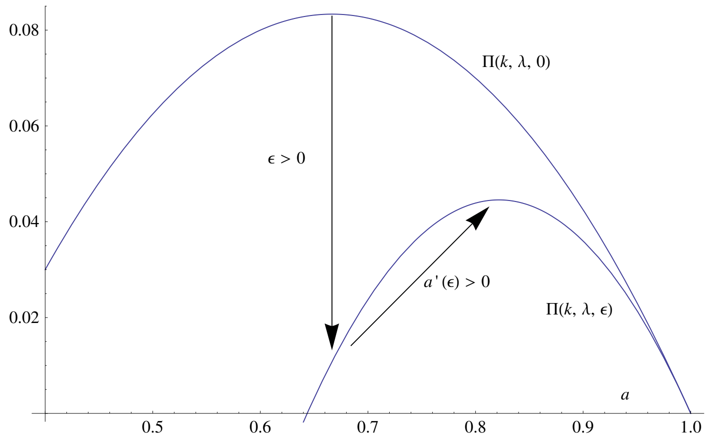

Figure 2: Dealer's profits as a function of ε

settlement failure for some k. Indeed, we have

$$\frac{\partial b(k)}{\partial \lambda_\varepsilon} = \frac{(1-\lambda)}{2\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left(1-\lambda-2k\lambda_\varepsilon-\lambda_\varepsilon^2\right)$$

and as Figure 3 shows, the bid price of dealer k increases with settlement risk if and only if:

$$k < \frac{1-\lambda-\lambda_\varepsilon^2}{2\lambda_\varepsilon} \equiv \kappa(\varepsilon). \tag{10}$$

Notice that $\kappa(\varepsilon) = 0$ whenever $\varepsilon = \sqrt{1-\lambda}\left(1-\sqrt{1-\lambda}\right)$. In general, one can easily

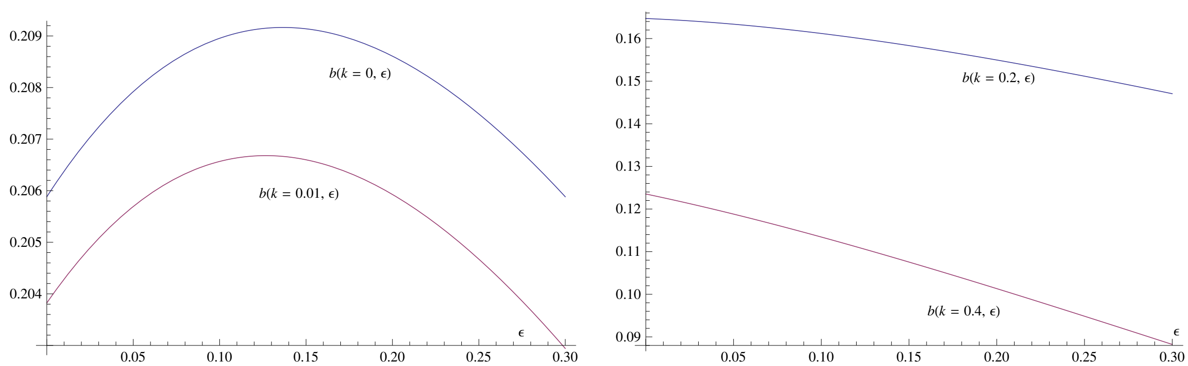

Figure 3: Bid prices as a function of ε for different dealers

prove the following result.

Lemma 1. *For all* $\varepsilon \leq \bar{\varepsilon} \equiv \sqrt{1-\lambda}\left(1-\sqrt{1-\lambda}\right)$, $b(k)$ *is increasing in* ε *whenever* $k < \kappa(\varepsilon)$ *and decreasing otherwise. For all* $\varepsilon > \bar{\varepsilon}$ *the bid price is always decreasing in* ε *for all* $k \leq \bar{k}$.

We can now characterize the demand and supply for each dealer:

$$\begin{aligned} D(a) &= \tfrac{1}{N}(1-a) = \frac{1}{2N}\frac{1-\lambda-k\lambda_\varepsilon}{(1-\lambda+\lambda_\varepsilon^2)} && (11)\\ S(b) &= \tfrac{1}{N}\lambda_\varepsilon(1-a) = \frac{1}{2N}\lambda_\varepsilon\frac{1-\lambda-k\lambda_\varepsilon}{(1-\lambda+\lambda_\varepsilon^2)} && (12)\end{aligned}$$

Substituting out for $a(k)$ and $b(k)$ from(8) and (9), as well as $N = \bar{k}$ in the profit function of dealer k, we obtain:

$$\Pi(k;\lambda,\varepsilon) = \frac{\lambda_\varepsilon(1-\lambda-k\lambda_\varepsilon)^2}{4(1-\lambda)\,(1-\lambda+\lambda_\varepsilon^2)} \tag{13}$$

Finally, the marginal active dealer $\bar{k}$ is such that $\Pi(\bar{k};\lambda,\varepsilon) = 0$, which yields:

$$\bar{k} = \frac{1-\lambda}{\lambda_\varepsilon} < 1. \tag{14}$$

It is then easy to see that $a\left(\bar{k}\right) = 1$. In the sequel, we show the main result of this section.

Lemma 2. *The dealers' surplus is decreasing in settlement risk. However, the most efficient dealers always benefit from an increase in settlement risk if and only if such risk is sufficiently small.*

Proof. Appendix A.1 shows that dealers' surplus is simply

$$S_d(\varepsilon) = \int_0^{\bar{k}} \Pi(k;\lambda,\varepsilon)dk = \frac{1}{12}\frac{(1-\lambda)^2}{\left(1-\lambda+{\lambda_\varepsilon}^2\right)}$$

which is always decreasing in ε. To see that the most efficient dealers benefit from an increase in settlement risk, notice that (13) implies that the marginal profits for dealer $k=0$ are:

$$\frac{\partial\Pi(0;\lambda,\varepsilon)}{\partial\varepsilon} = \frac{(1-\lambda)}{[4\,(1-\lambda+\lambda_\varepsilon^2)]^2}\left\{1-\lambda-\lambda_\varepsilon^2\right\}$$

which is increasing in ε whenever ε is small enough. In fact, the sign of $\partial\Pi(0;\lambda,\varepsilon)/\partial\varepsilon$ is the sign of $1-\lambda-\lambda_\varepsilon^2$. Hence, for all ε such that $\varepsilon < \bar{\varepsilon} = \sqrt{1-\lambda}(1-\sqrt{1-\lambda})$ the profit of the most efficient dealer will be increasing. $\square$

The surplus of consumers now has to take into account that consumers may not obtain the good if they fail to settle. Therefore, their surplus is scaled down by the probability of being hit by a settlement fail, λ. In Appendix A.6 we show that:

$$\begin{aligned} S_c(\bar{k}) &= (1-\lambda)\int_{a(0)}^{1}\left[\int_{a(0)}^{a(\bar{k})\vee v}\frac{(v-a)}{a(\bar{k})-a(0)}da\right]dv \\ &= \frac{1}{6}(1-\lambda)(1-a(0))^2 \end{aligned}$$

where $a(0) = 1 - \frac{1-\lambda}{2(1-\lambda+\lambda_\varepsilon^2)}$. Hence, the consumers' surplus is strictly decreasing with ε.[11] the following Lemma formalizes this result.

Lemma 3. *The consumers' surplus is decreasing with settlement risk.*

Finally, using the results in Appendix A.6, we compute the surplus of producers, as

$$S_p(\bar{k}) = \int_{0}^{b(0)}\left[\int_{b(\bar{k})\wedge v}^{b(0)}\frac{(b-v)}{b(0)-b(\bar{k})}db\right]dv = \frac{b(0)^2}{6}$$

where $b(0) = \lambda_\varepsilon\frac{1-\lambda}{2(1-\lambda+\lambda_\varepsilon^2)}$. Recall that Lemma 1 implies $\frac{\partial b(0)}{\partial\lambda_\varepsilon} > 0$ for λ_ε small enough, and $\frac{\partial b(0)}{\partial\lambda_\varepsilon} < 0$ otherwise. Therefore, the surplus of producers is increasing when there is little settlement risk, while it is decreasing otherwise. The following Lemma formalizes this result.

Lemma 4. *The producers' surplus is increasing with settlement risk whenever ε is small and it is decreasing otherwise.*

We now analyze whether the surplus for the entire economy is increasing in settlement risk. Hence, we define aggregate surplus as $S_d(\bar{k}) + S_p(\bar{k}) + S_c(\bar{k})$. It is more convenient to operate a change of variable to compute the surplus of dealers. In Appendix A we show

[11] This can be simplified to $S_c(\bar{k}) = \frac{1}{6}\frac{(1-\lambda)^3}{4(1-\lambda+\lambda_\varepsilon^2)^2}$.

that $S_d(\bar{k}) = \frac{2(1-\lambda+\lambda_\varepsilon^2)^2}{3(1-\lambda)}(1-a(0))^3$. Therefore, using results from Appendix A.6, aggregate surplus is simply:

$$\begin{aligned} S_d(\bar{k}) + S_p(\bar{k}) + S_c(\bar{k}) &= \frac{2(1-\lambda+\lambda_\varepsilon^2)^2}{3(1-\lambda)}(1-a(0))^3 \\ &+ \frac{b(0)^2}{6} + \frac{1}{6}(1-\lambda)(1-a(0))^2 \end{aligned}$$

and using $b(k) = \lambda_\varepsilon(1-a(k))$ and simplifying, we obtain

$$S \equiv S_d(\bar{k}) + S_p(\bar{k}) + S_c(\bar{k}) = \frac{(1-\lambda)^2}{8(1-\lambda+\lambda_\varepsilon^2)}$$

which is strictly decreasing in ε.

We summarize these results in the following proposition.

Proposition 1. *The consumers' expected surplus is decreasing with settlement risk as measured by ε. The producers' surplus is increasing in ε if ε is small enough, and it is decreasing otherwise. Aggregate dealers' surplus is decreasing in ε. However, the most efficient dealers always benefit from an increase in settlement risk. The overall welfare as measured by the equally weighted sum of all expected surplus is decreasing in ε.*

To conclude this section, we should stress that while it is efficient to reduce risk as much as possible, this is detrimental to the most efficient dealers. Less risk implies that less efficient dealers can profitably enter the market, thus making the market tighter for the most efficient dealers. In the next section, we analyze how these results affect dealers' decision to adopt a better market making technology.

5 Model with dealers' ex-ante fixed investment

In this section we study whether dealers have incentives to invest ex-ante into a technology that allows them to be more efficient in intermediating transactions between consumers and producers. Specifically, we assume that if dealers pay an effort cost γ then they draw their trading cost from a distribution which places larger probability on more efficient values of the support.

Because we interpret the trading cost k as a technology to intermediate transactions between consumers and producers, we refer to dealers' decision to exert effort as dealers' investment in the low cost technology. If, on the other hand, dealers do not exert effort then they draw their trading cost from a distribution with truncated support from the bottom. We refer to dealers' decision to not exert effort as dealers not investing, or adopting the high cost technology.

Intuitively, because a dealer is more efficient the lower its trading cost k and more efficient dealers earn larger profits from both a larger bid ask spread and from larger volume of intermediated transactions, then dealers have an incentive to invest in the more efficient technology as long as the cost γ is not too large. Because both consumers and producers benefit from being matched with more efficient dealers, dealers ex-ante investment also has benefits on the economy as a whole.[12] The introduction of a CCP, or of an interdealer market, however, by allowing more dealers to be profitable for a given level of counterparty risk (ε) may have the unintended consequence of reducing dealers' incentive to invest in the low cost technology, as more efficient dealers lose from the entry of relatively less efficient dealers who reduce their market share. When that happens, consumers and producers may also be worse off because they are less likely to be matched with efficient dealers and to trade.

5.1 Dealers' incentives to invest

We modify the benchmark model of the previous sections simply by adding an ex-ante choice for dealers. Because we want to maintain the tractable characteristics of the model developed in the previous sections, we maintain the assumption of uniform distribution of dealers' trading costs. We model dealers' choice as follows: if dealers invest ex ante by paying γ then they draw their trading cost from a uniform distribution on $[0, 1]$, which is the benchmark model analyzed in the previous sections. If dealers do not pay γ then they draw their trading cost from a uniform distribution on $[k_m, 1]$, with $k_m > 0$. Therefore the benchmark model represents the economy with the low cost technology, whereas the

[12]The fact that dealers do not necessarily rip those benefits turns out to not be crucial, since their incentives to invest in the low cost distribution is preserved under some assumptions.

characterization we derive below denotes the equilibrium in the economy with the high cost technology.

As in the benchmark model the marginal active dealer is the one who makes zero profits. We let $k_M \leq 1$ denote the type of such dealer, and

$$N = k_M - k_m$$

denote the measure of active dealers. As in the benchmark model, $D(a), S(b)$ denote the demand and supply of assets for each dealer when he posts ask price a and bid price b, $\Pi(k; \lambda, \varepsilon)$ denote the profits for a dealer with trading cost k and idiosyncratic risk $\varepsilon \in (0, \lambda)$ when consumers exit the economy with probability λ. Thus, with the measure of active dealers possibly different from the one in the benchmark model, we have

$$\begin{aligned} \Pi(k; \lambda, \varepsilon) &= \frac{1}{N} \frac{(1 - \lambda - k\lambda_\varepsilon)^2}{4(1 - \lambda + \lambda_\varepsilon^2)} \\ k_M &= \{k \in (k_m, 1) : \Pi(k; \lambda, \varepsilon) = 0\} \end{aligned}$$

and $k_m > 0$ given.

Lemma 5. $k_M = \overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$.

Proof. It follows from the derivation of $\overline{k}$ in the benchmark model (14) where N is replaced by $k_M - k_m$ rather than by $\overline{k}$. □

The surplus of dealers before they draw their type from a distribution $[k_m, 1]$ is the conditional expectation of their profits given by

$$\begin{aligned} S_d(\varepsilon; k_m) &= \int_{k_m}^{\overline{k}} \Pi(k; \lambda, \varepsilon) \frac{dk}{1 - k_m} \\ &= \frac{\left\{(1-\lambda)^2 N - \lambda_\varepsilon (1 - \lambda)\left(\overline{k} + k_m\right) N + \frac{\lambda_\varepsilon^2}{3}\left(\overline{k}^3 - k_m^3\right)\right\}}{N(1 - k_m) 4 (1 - \lambda + \lambda_\varepsilon^2)} \end{aligned} \tag{15}$$

In equation (15) the relevant distribution of dealers' transaction costs has been substituted out. When dealers do not invest in the low cost technology then they draw their k

from a uniform distribution over the support $[k_m, 1]$, with $k_m > 0$. Therefore, the probability that each dealer draws a specific $k \in [k_m, 1]$ is simply $\frac{1}{1-k_m}$. In other words, the distribution of dealers' transaction costs is truncated at $k_m > 0$. As a consequence, dealers' expected surplus ex-ante (i.e. before they draw their k) is the integral of a dealer's k profit over that probability measure. Similarly, with insurance against the idiosyncratic risk (recall $\lambda_\varepsilon = 1 - \lambda + \varepsilon$):

$$S_d\left(0; k_m\right) = \frac{1-\lambda}{4N\left(1-k_m\right)\left(2-\lambda\right)} \left\{ N\left(1 - \overline{k} - k_m\right) + \frac{\left(\overline{k}^3 - k_m^3\right)}{3} \right\} \tag{16}$$

Given some $\varepsilon > 0$ dealers have an incentive to invest in the low cost technology if and only if the ex-ante payoff from the investment in the low cost technology for a given idiosyncratic risk $\varepsilon > 0$, $S_d^L(\varepsilon) = S_d(\varepsilon; 0)$, net of the effort cost, exceeds the ex-ante payoff from not investing $S_d^H(\varepsilon) = S_d(\varepsilon; k_m)$, with $k_m > 0$, and drawing the trade cost from the high cost technology,

$$S_d^L(\varepsilon) - \gamma > S_d^H(\varepsilon)$$

Similarly, with full insurance against idiosyncratic risk, dealers lose the incentive to invest in the low cost distribution if and only if

$$S_d^L(0) - \gamma < S_d^H(0)$$

where, similarly to the case where $\varepsilon > 0$, $S_d^L(0) = S_d(0; 0)$ denotes the dealers' ex-ante surplus from investing in the low cost technology in an economy with no idiosyncratic risk, and where $S_d^H(0) = S_d(0; k_m)$, with $k_m > 0$, denotes the surplus from not investing and drawing the trade cost from the high cost technology in the same economy with no idiosyncratic risk.

Then dealers invest in the technology when there is risk iff the investment cost is small enough, but they do not invest when there is no risk iff the investment cost is too large. The following proposition shows that these bounds characterize a well defined and non-empty set of economies.

Proposition 2. *Given* $\varepsilon \in (0, \lambda]$, *assume* $k_m \in \left(0, \hat{k}\right)$ *with*

$$\hat{k} \;=\; \left[\frac{2\,(2-\lambda)\,\lambda_\varepsilon - (3-\lambda)\,(1-\lambda) - \lambda_\varepsilon^2}{\lambda_\varepsilon^2 - (1-\lambda)^2}\right](1-\lambda) \tag{17}$$

Then $S_d^L(\varepsilon) - S_d^H(\varepsilon) > \gamma > S_d^L(0) - S_d^H(0)$ *if and only if*

$$\overline{\gamma}_1(k_m, \varepsilon) \equiv \;\; \tfrac{(1-\lambda)(2\lambda_\varepsilon-(1-\lambda))k_m-(\lambda_\varepsilon k_m)^2}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} > \;\; \gamma > \frac{(1-\lambda)}{12\,(2-\lambda)}k_m \tag{18}$$

Proof. Consider first $S_d^L(\varepsilon) - S_d^H(\varepsilon) > \gamma$. Substituting out the equilibrium condition $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ yields

$$\overline{\gamma}_1(k_m, \varepsilon) = \frac{(1-\lambda)\,k_m\,(2\lambda_\varepsilon - (1-\lambda)) - (\lambda_\varepsilon k_m)^2}{12\,(1-k_m)\,(1-\lambda+\lambda_\varepsilon^2)} \;\; > \;\; \gamma \tag{19}$$

Consider now $\gamma > S_d^L(0) - S_d^H(0)$. Substituting out the equilibrium $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ yields

$$\gamma \;\; > \;\; \frac{(1-\lambda)}{12\,(2-\lambda)}k_m \tag{20}$$

Thus, a necessary condition for (19) and (20) to be satisfied is

$$\frac{(1-\lambda)\,(2\lambda_\varepsilon - (1-\lambda))\,k_m - (\lambda_\varepsilon k_m)^2}{12\,(1-k_m)\,(1-\lambda+\lambda_\varepsilon^2)} \;\; > \;\; \frac{(1-\lambda)}{12\,(2-\lambda)}k_m \tag{21}$$

which can be rearranged as

$$k_m \;\; < \;\; \left[\frac{2\,(2-\lambda)\,\lambda_\varepsilon - (3-\lambda)\,(1-\lambda) - \lambda_\varepsilon^2}{\lambda_\varepsilon^2 - (1-\lambda)^2}\right](1-\lambda)$$

and, more compactly, as $k_m < \hat{k}$, with $\hat{k}$ defined in (17).[13] □

Equation (17) defines an upper bound on k_m such that there exists a non degenerate set of economies indexed by $\gamma > 0$ – including the equilibrium described in proposition 2 – in which

[13] See Appendix B for full derivation of (19), (20) and (21).

dealers invest in the low cost technology in equilibrium if and only if they are not insured against idiosyncratic risk. That is to say that there exist economies such that condition (18) is satisfied. In Appendix 8, we show that such upper bound is never a binding constraint. In particular, it shows that in economies without insurance (i.e. $\varepsilon > 0$) the relevant upper bound on k_m for the assumptions in proposition 2 to be satisfied is $\overline{k}_\varepsilon = \frac{1-\lambda}{\lambda_\varepsilon}$, while in economies with insurance (i.e. $\varepsilon = 0$) it is $\hat{k}$ defined in (17). Thus, because the distribution of active dealers is $\mathcal{U}[k_m, \overline{k}_\varepsilon]$ in economies without insurance, then $k_m < \overline{k}_\varepsilon < \hat{k} < 1$. Furthermore, because our goal is to compare equilibrium outcomes in two economies which differ only with respect to insurance against idiosyncratic risk, then the assumption $k_m < \hat{k}$ in proposition 2 is always satisfied.

5.2 Equilibrium

An equilibrium is defined as in the benchmark model, except that dealers now have an additional decision to make. Before they draw their trading cost k they choose whether to incur a fixed cost of investing in the low-cost technology, for a given ε. If they do, then they pay a fixed effort cost γ and draw their k from a uniform distribution over $[0, 1]$, if they do not, then they draw their k from a uniform distribution over $[k_m, 1]$, with $k_m > 0$.

In the previous section we characterized the set of economies where an equilibrium is such that dealers prefer to invest in the low-cost technology if and only if they are not insured against idiosyncratic risk. These economies are characterized by intermediate values of the investment cost γ, as defined by condition (18). The investment cost needs to be sufficiently small to induce dealers to make the investment when they face idiosyncratic risk, but not too small so that dealers would still prefer to save on the effort cost when they are insured against idiosyncratic risk.

Moreover, we showed that if $k_m < \hat{k}$ then there always exists $\gamma > 0$ such that the conditions in proposition 2 are satisfied. Finally, lemma 8 in the Appendix implies that $\hat{k} > \overline{k}_\varepsilon$. Therefore there exists a non degenerate set of economies, indexed by $\gamma > 0$, such that the conditions in proposition 2 are satisfied. The following proposition formalizes results about existence and uniqueness of the equilibrium in these economies.

Proposition 3. *Let* $\overline{\gamma}_1(k_m, \varepsilon)$ *defined in (19) and assume* $\overline{\gamma}_1(k_m, \varepsilon) > \gamma > \frac{(1-\lambda)}{12(2-\lambda)} k_m$ *for*

$k_m > 0$. *Then there exists a unique equilibrium such that dealers invest in the low cost technology if and only if* $\varepsilon > 0$.

Proof. Because $\overline{\gamma}_1(k_m, \varepsilon) > \gamma > \frac{(1-\lambda)}{12(2-\lambda)} k_m$ by assumption, then condition (18) in proposition 2 are satisfied, implying that $S_d^L(\varepsilon) - S_d^H(\varepsilon) > \gamma > S_d^L(0) - S_d^H(0)$. Thus dealers invest in the low cost technology if and only if $\varepsilon > 0$. Existence and uniqueness of the equilibrium follow from the same arguments as in the benchmark model of the previous section. $\square$

5.3 Social planner's investment choice

Consider now the decision problem of a social planner who is constrained by the market mechanism[14] but can choose whether to pay the cost γ to invest in the technology that draws dealers' trading cost k from the distribution $\mathcal{U}[0,1]$ rather than the distribution $\mathcal{U}[k_m, 1]$. Because dealers are the agents who can invest in the low cost technology, then the planner is essentially choosing whether dealers should pay γ or not. In what follows we are agnostic about the issue of designing transfers that compensate dealers for their effort when the solution to the planner's problem involves paying γ.

The social planner maximizes ex-ante welfare of each type of agent, equally weighted. Thus, for a given $\varepsilon \geq 0$ the social planner chooses to pay γ if and only if

$$\sum_{j=d,c,p} \left[S_j^L(\varepsilon) - S_j^H(\varepsilon)\right] \;\; > \;\; \gamma.$$

In the previous section we showed conditions under which dealers choose to pay γ when $\varepsilon > 0$ but do not when $\varepsilon = 0$. Intuitively, both consumers and producers benefit from dealers' investment in drawing from the low cost technology, as they are matched with more efficient dealers and less often with less efficient dealers. Because a dealer's efficiency maps into her bid-ask spread and because more efficient dealers charge smaller bid-ask spreads, then both consumers and producers gain by dealers being more efficient on average. When $\varepsilon > 0$ and k_m satisfies (17), condition (19) implies that the increase in dealers' surplus from investing is sufficient to compensate them for paying γ. Then it is easy to show that the

[14] That is the social planner is subject to dealers having to intermediate transactions between consumers and producers, as they are permanently separated from each other.

social planner's solution also involves paying γ. When $\varepsilon = 0$, however, the social planner's allocation involves paying γ if and only if the resulting surplus of consumers and producers more than compensate the decrease in dealers' surplus net of γ, or:

$$S_c^L(0) - S_c^H(0) + S_p^L(0) - S_p^H(0) \quad > \quad \gamma - \left(S_d^L(0) - S_d^H(0)\right)$$

Because the low cost technology draws dealers from $\mathcal{U}[0,1]$ then $S_c^L(0)$ and $S_p^L(0)$ are the same as in the benchmark model. The characterization of the surplus of consumers and producers under the high cost technology for dealers requires a few additional steps, which are described below.

5.3.1 Consumers' surplus

Here we show that consumers always benefit from dealers investing in the most efficient technology. Consider consumers' surplus first, for a given $\varepsilon \geq 0$:

$$S_c(\underline{a}, \overline{a}; \varepsilon) \quad = \quad \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\int_{\underline{a}}^{\overline{a}}\int_{\underline{a}}^{v}(v-a)\,dadv + \int_{\overline{a}}^{1}\int_{\underline{a}}^{\overline{a}}(v-a)\,dadv\right]$$

for $\underline{a} = a(k_m)$ and $\overline{a} = a\left(\overline{k}\right)$ where the ask price function is characterized in (8)

$$a(k) \quad = \quad \frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2\left(1-\lambda+\lambda_\varepsilon^2\right)}$$

As in the benchmark model, because a dealer must purchase the asset from producers before selling it to consumers, and because a dealer must pay the trading cost for each asset purchased, then the lowest ask price is offered by the most efficient dealer and the largest by the least efficient dealer. Efficient dealers make higher profits per transaction than less efficient dealers, thus they can afford being paid a lower price per asset sold to a consumer than less efficient dealers. In the appendix we show that $S_c(\underline{a}, \overline{a}; \varepsilon)$ can be rewritten as

$$S_c(\underline{a}, \overline{a}; \varepsilon) \quad = \quad \frac{1-\lambda}{6}\left[3+(\overline{a}+\underline{a})(\overline{a}-3)+\underline{a}^2\right] \tag{22}$$

In the appendix we also show that the increase in the consumers' surplus from dealers' investment in the low cost technology is positive, as

$$\begin{aligned} S_c^L(\varepsilon) - S_c^H(\varepsilon) &= S_c(a(0), 1; \varepsilon) - S_c(\underline{a}, 1; \varepsilon) \\ &= \frac{(1-\lambda) k_m \lambda_\varepsilon}{24(1-\lambda+\lambda_\varepsilon^2)^2} [2(1-\lambda) - k_m \lambda_\varepsilon] > 0 \end{aligned} \tag{23}$$

where the last inequality follows from $2(1-\lambda) > k_m \lambda_\varepsilon$ since $\varepsilon \in (0, \lambda)$ and $k_m < \frac{1-\lambda}{\lambda_\varepsilon} = \overline{k}_\varepsilon$. To ease notation, as in the case of dealers' surplus, we define $S_c^H(\varepsilon) = S_c(\underline{a} = a(k_m), 1; \varepsilon)$ with $k_m > 0$ as the consumers' surplus when dealers do not invest in the low cost technology for a given pair $k_m > 0, \varepsilon > 0$. Analogously, $S_c^L(\varepsilon) = S_c(\underline{a} = a(0), 1; \varepsilon)$ is the consumers' surplus when dealers invest in the low cost technology, resulting in $k_m = 0$, for a given $\varepsilon > 0$. In particular, in the case where $\varepsilon = 0$, the increase in consumers' surplus is:

$$S_c^L(0) - S_c^H(0) = \frac{k_m(1-\lambda)(2-k_m)}{24(2-\lambda)^2} \tag{24}$$

Therefore, combining (23) with (24) we conclude that consumers always benefit from the investment in the low cost technology, for all $\varepsilon \geq 0$.

5.3.2 Producers' surplus

In this section we show that producers always benefit from the investment in the low cost technology for all $\varepsilon \geq 0$. For any $\varepsilon \geq 0$, the producers' surplus is

$$S_p\left(\underline{b}, \overline{b}; \varepsilon\right) = \frac{1}{\left(\overline{b} - \underline{b}\right)} \left[\int_{\underline{b}}^{\overline{b}} \int_{v}^{\overline{b}} (b - v)\, dbdv + \int_0^{\underline{b}} \int_{\underline{b}}^{\overline{b}} (b - v)\, dbdv\right]$$

for $\underline{b} = b\left(\overline{k}_\varepsilon\right)$ and $\overline{b} = b(k_m)$ where the bid price function is characterized in (9)

$$b(k) = = \lambda_\varepsilon (1 - a(k)) = \lambda_\varepsilon \frac{(1-\lambda-k\lambda_\varepsilon)}{2(1-\lambda+\lambda_\varepsilon^2)}$$

Again, as in the benchmark model, the lowest bid price is offered by the least efficient dealer and the highest by the most efficient dealer. Efficient dealers make higher profits per transaction than less efficient dealers, thus they can afford to pay a higher price per asset purchased from a producer than less efficient dealers. In the appendix we show that $S_p\left(\underline{b},\overline{b};\varepsilon\right)$ can be rewritten as

$$S_p\left(\underline{b},\overline{b};\varepsilon\right) = \frac{\overline{b}^2+\overline{b}\underline{b}+\underline{b}^2}{6}$$

Notice that $\underline{b}=b\left(\overline{k}_\varepsilon\right)=0$, since even inefficient dealers can afford to purchase the asset owned by a producer with valuation for the assets $v=0$. On the other hand $\overline{b}=b\left(k_m\right)=\lambda_\varepsilon\frac{(1-\lambda-k_m\lambda_\varepsilon)}{2(1-\lambda+\lambda_\varepsilon^2)}$. Then, substituting out $\underline{b}=0$ in the producers' surplus yields $S_p\left(0,\overline{b};\varepsilon\right)=\frac{\overline{b}^2}{6}$, with $\overline{b}=b\left(k_m\right)$. Because $\underline{b}=0$ regardless of the distribution dealers are drawn from, then also $S_p^L\left(0,\overline{b};\varepsilon\right)=\frac{\overline{b}^2}{6}$, with $\overline{b}=b\left(0\right)$. The gain in the surplus of producers from dealers' investment into the low cost technology is then:

$$S_p\left(0,b\left(0\right);\varepsilon\right)-S_p\left(0,b\left(k_m\right);\varepsilon\right) = \frac{b\left(0\right)^2-b\left(k_m\right)^2}{6}.$$

In the appendix we show that

$$S_p\left(0,b\left(0\right);\varepsilon\right)-S_p\left(0,b\left(k_m\right);\varepsilon\right) = \frac{k_m\lambda_\varepsilon^3\left[2\left(1-\lambda\right)-k_m\lambda_\varepsilon\right]}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2} \tag{25}$$

where, by feasibility, $k_m<\overline{k}=\frac{1-\lambda}{\lambda_\varepsilon}$. Thus

$$S_p\left(0,b\left(0\right);\varepsilon\right)-S_p\left(0,b\left(k_m\right);\varepsilon\right) > \frac{k_m\lambda_\varepsilon^3\left[2\left(1-\lambda\right)-\frac{1-\lambda}{\lambda_\varepsilon}\lambda_\varepsilon\right]}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}>0 \tag{26}$$

To ease notation as in the previous cases of dealers' and consumers' surpluses, we define $S_p^L\left(\varepsilon\right)=S_p\left(0,b\left(0\right);\varepsilon\right)$ as the producers' surplus when dealers invest in the low cost technology, resulting in $k_m=0$, for a given $\varepsilon>0$. Analogously, we define $S_p^H\left(\varepsilon\right)=S_p\left(0,b\left(k_m\right);\varepsilon\right)$ as the producers' surplus when dealers do not invest in the low cost technology, for a given pair $k_m > 0, \varepsilon > 0$. Equation (25) implies that when $\varepsilon = 0$ the gain in producers' surplus is

$$S_p^L(0) - S_p^H(0) = \frac{k_m (1-\lambda)^2 (2-k_m)}{24(2-\lambda)^2} \tag{27}$$

Combining (26) with (27) we conclude that producers always benefit from the investment in the low cost technology for all $\varepsilon \geq 0$.

5.3.3 Social planner's solution

We now analyze the decision problem of a social planner who can choose whether to force dealers to invest in the low cost technology. We consider two sets of economies: one with no idiosyncratic risk (i.e. $\varepsilon = 0$) and one with idiosyncratic risk (i.e. $\varepsilon > 0$). The aim of this exercise is to check the efficiency of the equilibrium characterized in the previous section, where dealers do not invest in the low cost technology when their idiosyncratic risk is insured. We will perform the same exercise for an economy with idiosyncratic risk (i.e. $\varepsilon > 0$). Intuitively, in these economies, if dealers invest in the low cost technology in equilibrium, then it must be efficient. In fact, for investment to be part of an equilibrium, it must be that the cost of investing in the low cost technology is lower than dealers' net gain from such investment ($\gamma < S_d^L(\varepsilon) - S_d^H(\varepsilon)$). Because (23) and (26) imply that the gain in consumers' and producers' surpluses from dealers' investment is always positive for all $\varepsilon > 0$, a social planner would also choose to invest in the low cost technology. Thus the equilibrium is efficient. However, in equilibrium dealers do not account for the effects of their investment decision on the surplus of consumers and producers. So there may be a set of economies where γ is too large for dealers to invest in the technology, while still low enough for the social planner to prefer investing. Then, these economies will be inefficient.

Consider first an economy with no idiosyncratic risk (i.e. $\varepsilon = 0$). The solution to the social planner's problem is to invest if and only if the sum of the gains in consumers' and producers' surplus exceeds the loss in dealers' surplus net of the investment cost:

$$S_c^L(0) - S_c^H(0) + S_p^L(0) - S_p^H(0) > \gamma - \left[S_d^L(0) - S_d^H(0)\right]$$

Using the characterizations of the gains in agents' surpluses derived in the previous sections, this inequality simplifies to:

$$\overline{\gamma}_2\left(k_m,0\right)\equiv\frac{k_m\left(1-\lambda\right)\left(4-k_m\right)}{24\left(2-\lambda\right)}\quad>\quad\gamma \tag{28}$$

$\overline{\gamma}_2\left(k_m,0\right)$ sets an upper bound on γ, which is increasing in k_m: since for higher values of k_m the gains from adopting the better technology are higher for all agents, then the planner is willing to pay a higher price for it.[15] For all economies such that γ is too large for dealers to be willing to invest (i.e. $\gamma>\overline{\gamma}_1\left(k_m,\varepsilon\right)$ as defined in (19)) but sufficiently small for the planner to invest (i.e. $\overline{\gamma}_2\left(k_m,0\right)\geq\gamma$), the equilibrium is inefficient. We summarize these results in the following proposition.

Proposition 4. *Consider economies where $\varepsilon=0$. As in (20), assume $\gamma>\frac{(1-\lambda)}{12(2-\lambda)}k_m$ and $k_m\in(0,\hat{k})$, with $\hat{k}$ defined in (17). The equilibrium is inefficient if and only if $\overline{\gamma}_2\left(k_m,0\right)\geq\gamma$.*

Proof. Inequality (28) defines the upper bound for γ such that the social planner chooses to invest. Since dealers prefer not to invest if $\gamma>\frac{(1-\lambda)}{12(2-\lambda)}k_m$, then an equilibrium where dealers are insured against idiosyncratic risk is inefficient iff

$$\overline{\gamma}_2\left(k_m,0\right)\quad>\gamma\quad>\frac{\left(1-\lambda\right)}{12\left(2-\lambda\right)}k_m \tag{29}$$

Necessary and sufficient condition (29) can be rearranged as:

$$k_m\frac{\left(1-\lambda\right)\left(4-k_m\right)}{24\left(2-\lambda\right)}>\gamma\quad>\quad\frac{\left(1-\lambda\right)}{12\left(2-\lambda\right)}k_m$$

A necessary condition for the existence of $\gamma>0$ such that the above inequality is satisfied is $k_m\frac{(1-\lambda)(4-k_m)}{24(2-\lambda)}>\frac{(1-\lambda)}{12(2-\lambda)}k_m$, which is always satisfied since $k_m<\overline{k}<1$. □

Consider now economies with idiosyncratic risk (i.e. $\varepsilon>0$). The solution to the social planner's problem is to invest if and only if the sum of the gains in consumers' and producers'

[15] This upper bound must be consistent with the upper bound set by (19) for an equilibrium to be also such that dealers invest in the low cost technology when $\varepsilon>0$, rather than having dealers never willing to invest in equilibrium. This is simply to have a trade off between insurance and incentives to invest in equilibrium.

surplus exceeds the loss in dealers' surplus net of the investment cost:

$$S_c^L(\varepsilon) - S_c^H(\varepsilon) + S_p^L(\varepsilon) - S_p^H(\varepsilon) + S_d^L(\varepsilon) - S_d^H(\varepsilon) \quad > \quad \gamma$$

In the appendix we show that this can be rearranged as:

$$\begin{aligned} \overline{\gamma}_2(k_m, \varepsilon) \equiv \frac{(1-k_m)\, k_m \lambda_\varepsilon \left[2(1-\lambda) - k_m \lambda_\varepsilon\right]}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} & \quad + \\ \frac{(1-\lambda)\, k_m 4\lambda_\varepsilon - 2(1-\lambda)^2 k_m - 2(\lambda_\varepsilon k_m)^2}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} & \quad > \quad \gamma \end{aligned} \tag{30}$$

Then the following proposition characterizes necessary and sufficient conditions for investment in the low cost technology to solve the social planner's problem.

Proposition 5. *Consider economies where $\varepsilon > 0$. The solution to the social planner's problem is to invest in the low cost technology if and only if $\overline{\gamma}_2(k_m, \varepsilon) \geq \gamma$.*

Proof. It follows from (30). □

Now we can compare $\overline{\gamma}_2(k_m, \varepsilon)$ with the relevant threshold of γ for dealers to invest in the low cost technology, $\overline{\gamma}_1(k_m, \varepsilon)$, defined in (19). Intuitively, the threshold of γ defining the maximum effort cost for dealers such that the social planner invests in the low cost technology should be larger than the threshold of γ above which dealers no longer invest in the low cost technology. In fact, in the previous sections we showed that the gain in both consumers' and producers' surplus from the investment is always strictly positive for all $\varepsilon \geq 0$. Because the gain in both consumers' and producers' surplus is relevant for the decision of the social planner but not for the decision of dealers individually, then it must be that the maximum effort cost γ such that the social planner invests in the low cost technology, $\overline{\gamma}_2(k_m, \varepsilon)$ is larger than the maximum effort cost such that dealers invest in the low cost technology, $\overline{\gamma}_1(k_m, \varepsilon)$. The following lemma formalizes this intuition.

Lemma 6. $\overline{\gamma}_2(k_m, \varepsilon) > \overline{\gamma}_1(k_m, \varepsilon)$ *for all* $\lambda \in (0, 1)$, $\varepsilon \in (0, \lambda)$.

Proof. The left hand side of (30) defines $\overline{\gamma}_2(k_m, \varepsilon)$ and (19) defines $\overline{\gamma}_1(k_m, \varepsilon)$. So $\overline{\gamma}_2(k_m, \varepsilon) >$

$\overline{\gamma}_1(k_m, \varepsilon)$ if and only if:

$$\frac{(1-k_m)\,k_m\lambda_\varepsilon\left[2(1-\lambda)-k_m\lambda_\varepsilon\right]}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)}+\frac{(1-\lambda)\,k_m 4\lambda_\varepsilon-2(1-\lambda)^2 k_m-2(\lambda_\varepsilon k_m)^2}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} > \frac{(1-\lambda)(2\lambda_\varepsilon-(1-\lambda))\,k_m-(\lambda_\varepsilon k_m)^2}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)}$$

which can be rearranged as

$$\frac{(1-k_m)\,k_m\lambda_\varepsilon\left[2(1-\lambda)-k_m\lambda_\varepsilon\right]}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} > 0$$

and is always satisfied because $[2(1-\lambda)-k_m\lambda_\varepsilon] > 0$. □

Finally, we can conclude this section with our main result, which is merely a corollary to Proposition 6.

Corollary 1. *Consider economies where $\varepsilon > 0$. This equilibrium is inefficient if and only if $\overline{\gamma}_2(k_m, \varepsilon) > \gamma > \overline{\gamma}_1(k_m, \varepsilon)$.*

Proof. In equilibrium dealers do not invest in the low cost technology, because $\gamma > \overline{\gamma}_1(k_m, \varepsilon)$. However, because $\overline{\gamma}_2(k_m, \varepsilon) > \gamma$, the solution to the social planner is to invest. Therefore the equilibrium is inefficient. Conversely, if the equilibrium is inefficient, it must be that the social planner chooses to invest, which is the case if and only if $\overline{\gamma}_2(k_m, \varepsilon) > \gamma$, as shown in the proposition 5. □

Corollary 1 states conditions under which the economy with risk is inefficient, because dealers prefer to keep an inefficient market making technology while the planner would rather have them invest in a better one. Finally, let us stress that Proposition 3 implies that the economy could be efficient for $\varepsilon > 0$ but inefficient for $\varepsilon = 0$, so that reducing risk can make a representative investor worse off.

5.4 Average bid-ask spreads

Consider economies where the assumptions of Proposition 3 are satisfied. This guarantees that, in equilibrium, dealers invest in the low cost technology if and only if they face some risk (i.e. $\varepsilon > 0$). With insurance (i.e. $\varepsilon = 0$) dealers do not invest in the low cost technology. This has consequences for the equilibrium average bid-ask spread observed in the market where dealers intermediate transactions between buyers and sellers.

In this section we show that, due to the general equilibrium effect of insurance on dealers' incentives to invest in ex-ante efficient technologies, the impact of central clearing on average bid-ask spreads is ambiguous and depends on the ex-ante characteristics of dealers.[16] In particular, comparing the economy with insurance to the economy without, we are able to characterize a necessary and sufficient condition on dealers' distribution of transaction costs for the bid-ask spread to be smaller in the economy with insurance. This requires the minimum transaction cost for dealers (k_m) to be sufficiently small. Intuitively, insurance causes bid-ask spreads to shrink which, in turn, fosters competition by allowing less efficient dealers to enter the market and be profitable. On the other hand, insurance has a perverse indirect effect on the incentives of dealers to invest ex-ante in a more efficient technology. As the ex-ante pool of dealers becomes worse (from $[0, \overline{k}]$ to $[k_m, 1]$), the average bid-ask spread may increase in equilibrium, as a dealer's quoted bid-ask spread depends on its transaction cost, as implied by (8) and (9), with the bid-ask spread decreasing in the efficiency of a dealer (i.e. the most efficient dealer charges the smallest bid-ask spread). As a result, the general equilibrium effect of central clearing on bid-ask spreads is negative (i.e. central clearing is associated with smaller bid-ask spreads) if the first effect dominates. This happens when the pool of dealers does not become *too* worse when dealers stop investing in the ex-ante efficient technology.

The following proposition formalizes this results.

Proposition 6. *Maintain the assumptions in Proposition 3. Then average bid-ask spread is smaller is the economy with insurance and the high cost technology if and only if:*

$$\frac{\lambda_\varepsilon^2 - \lambda_\varepsilon (1 - \lambda)}{(1 - \lambda + \lambda_\varepsilon^2)} \quad > \quad k_m \tag{31}$$

Proof. See appendix. □

[16]See Appendix D for a characterization of the mapping between a reduction in settlement risk and the introduction of central clearing.

Proposition 6 is particularly relevant in light of recent empirical findings by Loon and Zhong [2016] who study the effect of central clearing on a measure of transaction-level spreads. They analyze individually the three phases of mandatory central clearing implementation by the CFTC, with each phase covering a different category of market participants. Loon and Zhong [2016] find that central clearing is associated with an increase in spreads for swap dealers (Phase 1), while it is associated with a decrease in spreads for commodity pools (Phase 2) and all other swap market participants (Phase 3).[17] Our results suggest that differences in dealers' ex-ante characteristics, such as the support of the distribution of trading costs, are responsible for differences in bid-ask spreads as the provision of insurance via central clearing affects equilibrium prices directly and indirectly in opposite directions.

6 Conclusion

Market makers are useful to solve several frictions prevalent in financial markets. In this paper we concentrated on the effect of search frictions. The presence of frictions in general implies that market makers can earn a rent. Not surprisingly, this rent is proportional to a dealer's efficiency in making market. The more efficient a dealer is the higher his rent. However, this rent may be declining in the working efficiency of markets. For example, introducing an insurance against inventory risk (which arises from settlement risk in our model) can reduce the rent of the most efficient dealers because less efficient dealers can now operate thus increasing competition. While this looks like a desirable outcome, we show that this can be detrimental to welfare whenever the decision to be "more efficient" is endogenous. By lowering the benefit of being better at making markets, technological innovations in the structure of market can induce market makers to stop investing in better market making technologies, thus hampering the effects of the innovations. The paper thus offers a perspective on the opposition of some dealers to the recent pressure for improving market structures, such as clearing all derivatives traded OTC on central counterparties. Second, it argues that forcing the adoption of seemingly better market infrastructure has consequences for the incentives of some market participants, which can adversely impact

[17] See Loon and Zhong [2016], Appendix A.2.1, pg. 669.

other agents. Controlling for these incentives, possibly through transfers, is key to rip the entire gains from the better market structure.

References

FSB. Review of OTC derivatives market reforms. Effectiveness and broader effects of the reforms. Technical report, Financial Stability Board, 2017.

Yee Cheng Loon and Zhaodong (Ken) Zhong. Does Dodd-Frank affect OTC transaction costs and liquidity? Evidence from real-time CDS trade reports. *Journal of Financial Economics*, 119:645–672, 2016.

Darrell Duffie, Nicolae Gârleanu, and Lasse Heje Pedersen. Over-the-Counter Markets. *Econometrica*, 73(6):1815–1847, 2005.

William C. Dudley. Reforming the otc derivatives market. In *Remarks at the Harvard Law School's Symposium on Building the Financial System of the 21st Century, Armonk, New York*. http://www.newyorkfed.org/newsevents/speeches/2012/dud120322.html, March 22, 2012.

Maureen O'hara. *Market microstructure theory*, volume 108. Blackwell Cambridge, MA, 1995.

Yakov Amihud and Haim Mendelson. Dealership market: Market-making with inventory. *Journal of Financial Economics*, 8(1):31–53, 1980.

Jürgen Dennert. Price competition between market makers. *The Review of Economic Studies*, 60(3):735–751, 1993.

Tano Santos and Jose A Scheinkman. Competition among exchanges. *The Quarterly Journal of Economics*, 116(3):1027–1061, 2001.

Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

D. Spulber. Market making by price-setting firms. *Review of Economic Studies*, 63:559–580, 1996.

John Rust and George Hall. Middlemen versus Market Makers: A Theory of Competitive Exchange. *The Journal of Political Economy*, 111:2:353–403, 2003.

Pierre-Olivier Weill. Leaning against the wind. *The Review of Economic Studies*, 74(4): 1329–1354, 2007.

Pierre-Olivier Weill. Liquidity provision in capacity-constrained markets. *Macroeconomic Dynamics*, 15(S1):119–144, 2011.

Andrew G Atkeson, Andrea L Eisfeldt, and Pierre-Olivier Weill. Entry and exit in otc derivatives markets. *Econometrica*, 83(6):2231–2292, 2015.

Bank for International Settlements BIS. A glossary of terms used in payments and settlement systems, `https://www.bis.org/publ/cpss00b.pdf`. Technical report.

ISDA. CCP Loss Allocation at the End of the Waterfall. Technical report, 2013.

Default Waterfall. EUREX clearing, 2017. URL `http://www.eurexclearing.com/clearing-en/risk-management/default-waterfall`.

How Clearing works. Inter Continental Exchange, 2017. URL `https://www.theice.com/publicdocs/How_Clearing_Works.pdf`.

DTCC Publishes Perspectives on CCP Risk Management . EUREX clearing, 2017. URL `http://www.dtcc.com/news/2015/june/01/dtcc-publishes-perspectives-on-ccp-risk-management`.

CCP Risk Management Recovery & Resolution. London Clearing House, 2017. URL `http://www.lch.com/documents/731485/762448/ccpriskmanagement_whitepaper.pdf/`.

Viral Acharya and Alberto Bisin. Counterparty risk externality: Centralized versus over-the-counter markets. *Journal of Economic Theory*, 149:153–182, 2014.

Thorsten V Koeppl and Cyril Monnet. Central counterparty clearing and systemic risk insurance in otc derivatives markets. *forthcoming in a special issue of the Revue d'Economie Financiere*, 2013.

Thorsten Koeppl, Cyril Monnet, and Ted Temzelides. Optimal clearing arrangements for financial trades. *Journal of Financial Economics*, 103(1):189–203, 2012.

Bruno Biais, Florian Heider, and Marie Hoerova. Risk-sharing or risk-taking? counterparty risk, incentives, and margins. *The Journal of Finance*, 71(4):1669–1698, 2016.

Bruno Biais, Florian Heider, and Marie Hoerova. Clearing, counterparty risk, and aggregate risk. *IMF Economic Review*, 60(2):193–222, 2012.

A Derivations

A.1 Derivation of $S_d(\varepsilon)$

$$\begin{aligned}
S_d(\varepsilon) &= \int_{k_m}^{\overline{k}} \Pi(k;\lambda,\varepsilon)\frac{dk}{N} \\
&= \frac{1}{N^2 4\left(1-\lambda+\lambda_\varepsilon^2\right)}\int_{k_m}^{\overline{k}}\left(1-\lambda-k\lambda_\varepsilon\right)^2 dk \\
&= \frac{1}{\left(\overline{k}-k_m\right)^2 4\left(1-\lambda+\lambda_\varepsilon^2\right)}\int_{k_m}^{\overline{k}}\left(1-\lambda-k\lambda_\varepsilon\right)^2 dk \\
&= \frac{1}{\left(\overline{k}-k_m\right)^2 4\left(1-\lambda+\lambda_\varepsilon^2\right)}\int_{k_m}^{\overline{k}}\left((1-\lambda)^2+\left(k\lambda_\varepsilon\right)^2-2(1-\lambda)k\lambda_\varepsilon\right) dk \\
&= \frac{1}{\left(\overline{k}-k_m\right)^2 4\left(1-\lambda+\lambda_\varepsilon^2\right)}\left\{(1-\lambda)^2\left(\overline{k}-k_m\right)-2(1-\lambda)\lambda_\varepsilon\frac{\overline{k}^2-k_m^2}{2}+\lambda_\varepsilon^2\frac{\overline{k}^3-k_m^3}{3}\right\} \\
&= \frac{1}{\left(\overline{k}-k_m\right)^2 4\left(1-\lambda+\lambda_\varepsilon^2\right)}\left\{(1-\lambda)^2 N-\lambda_\varepsilon(1-\lambda)\left(\overline{k}+k_m\right)N+\frac{\lambda_\varepsilon^2}{3}\left(\overline{k}^3-k_m^3\right)\right\} \qquad (32)
\end{aligned}$$

Thus

$$S_d(0) = \frac{1}{N^2 4(1-\lambda)(2-\lambda)}\left\{(1-\lambda)^2 N-\lambda_\varepsilon(1-\lambda)\left(\overline{k}+k_m\right)N+\frac{(1-\lambda)^2}{3}\left(\overline{k}^3-k_m^3\right)\right\}$$

$$= \quad \frac{1-\lambda}{4N^2\left(2-\lambda\right)}\left\{N\left(1-\overline{k}-k_m\right)+\frac{\left(\overline{k}^3-k_m^3\right)}{3}\right\} \tag{33}$$

A.2 Derivation of (19)

Consider $S_d^L(\varepsilon) - S_d(\varepsilon) > \gamma$. Equation (15) can be rearranged as:

$$\begin{aligned}
S_d(\varepsilon) &= \frac{(1-\lambda)\left[1-\lambda-\lambda_\varepsilon\left(\overline{k}+k_m\right)\right]}{4N\left(1-\lambda+\lambda_\varepsilon^2\right)} + \frac{\lambda_\varepsilon^2\left(\overline{k}^3-k_m^3\right)}{12N^2\left(1-\lambda+\lambda_\varepsilon^2\right)} \\
&= \frac{(1-\lambda)\left[1-\lambda-\lambda_\varepsilon\left(\frac{1-\lambda}{\lambda_\varepsilon}+k_m\right)\right]}{4\left(\frac{1-\lambda}{\lambda_\varepsilon}-k_m\right)\left(1-\lambda+\lambda_\varepsilon^2\right)} + \frac{\lambda_\varepsilon^2\left(\frac{(1-\lambda)^3}{\lambda_\varepsilon^3}-k_m^3\right)}{12\left(\frac{(1-\lambda)}{\lambda_\varepsilon}-k_m\right)^2\left(1-\lambda+\lambda_\varepsilon^2\right)} \\
&= -\frac{\lambda_\varepsilon k_m(1-\lambda)}{4\left(\frac{1-\lambda}{\lambda_\varepsilon}-k_m\right)\left(1-\lambda+\lambda_\varepsilon^2\right)} + \frac{\frac{(1-\lambda)^3}{\lambda_\varepsilon}-\lambda_\varepsilon^2k_m^3}{12\left(\frac{(1-\lambda)}{\lambda_\varepsilon}-k_m\right)^2\left(1-\lambda+\lambda_\varepsilon^2\right)} \\
&= \frac{1}{4\left(\frac{1-\lambda}{\lambda_\varepsilon}-k_m\right)\left(1-\lambda+\lambda_\varepsilon^2\right)}\left\{-\lambda_\varepsilon k_m(1-\lambda)+\frac{(1-\lambda)^3-\lambda_\varepsilon^3k_m^3}{3\lambda_\varepsilon\left(\frac{(1-\lambda)}{\lambda_\varepsilon}-k_m\right)}\right\} \\
&= \frac{\lambda_\varepsilon}{4\left(1-\lambda-\lambda_\varepsilon k_m\right)\left(1-\lambda+\lambda_\varepsilon^2\right)}\left\{\frac{(1-\lambda)^3-\lambda_\varepsilon^3k_m^3}{3\left(1-\lambda-\lambda_\varepsilon k_m\right)}-\lambda_\varepsilon k_m(1-\lambda)\right\} \qquad (34)
\end{aligned}$$

where $\overline{k}=\frac{1-\lambda}{\lambda_\varepsilon}$ has been substituted out.

Thus $S_d^L(\varepsilon) - S_d(\varepsilon) > \gamma$ if and only if

$$\frac{(1-\lambda)^2}{12\left(1-\lambda+\lambda_\varepsilon^2\right)} - \frac{\lambda_\varepsilon}{4\left(1-\lambda-\lambda_\varepsilon k_m\right)\left(1-\lambda+\lambda_\varepsilon^2\right)}\left\{\frac{(1-\lambda)^3-\left(\lambda_\varepsilon k_m\right)^3}{3\left(1-\lambda-\lambda_\varepsilon k_m\right)}-\lambda_\varepsilon k_m(1-\lambda)\right\} \quad > \quad \gamma$$

which can be rewritten as

$$\begin{aligned}
\frac{(1-\lambda)^2-\frac{\lambda_\varepsilon}{(1-\lambda-\lambda_\varepsilon k_m)}\left[\frac{(1-\lambda)^3-(\lambda_\varepsilon k_m)^3}{(1-\lambda-\lambda_\varepsilon k_m)}-3\lambda_\varepsilon k_m(1-\lambda)\right]}{12\left(1-\lambda+\lambda_\varepsilon^2\right)} &> \gamma \\
\frac{(1-\lambda)^2-\frac{\lambda_\varepsilon}{(1-\lambda-\lambda_\varepsilon k_m)^2}\left[(1-\lambda)^3-(\lambda_\varepsilon k_m)^3-3\lambda_\varepsilon k_m(1-\lambda)^2+3\lambda_\varepsilon k_m(1-\lambda)\lambda_\varepsilon k_m\right]}{12\left(1-\lambda+\lambda_\varepsilon^2\right)} &> \gamma
\end{aligned}$$

$$\frac{(1-\lambda)^2 - \frac{\lambda_\varepsilon(1-\lambda-\lambda_\varepsilon k_m)^3}{(1-\lambda-\lambda_\varepsilon k_m)^2}}{12\left(1-\lambda+\lambda_\varepsilon^2\right)} \quad > \quad \gamma$$

and which finally yields (19):

$$\frac{(1-\lambda)^2 - \lambda_\varepsilon\left(1-\lambda-\lambda_\varepsilon k_m\right)}{12\left(1-\lambda+\lambda_\varepsilon^2\right)} \quad > \quad \gamma$$

A.3 Derivation of (20)

Consider $\gamma > S_d^L(0) - S_d(0)$. Equation (15) can be rearranged using $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ as (34), which, evaluated at $\varepsilon = 0$ yields:

$$\begin{aligned}
S_d(0) &= \frac{1-\lambda}{4\left(\frac{1-\lambda}{\lambda_\varepsilon}-k_m\right)^2(2-\lambda)}\left\{\left(\frac{1-\lambda}{\lambda_\varepsilon}-k_m\right)\left(1-\frac{1-\lambda}{\lambda_\varepsilon}-k_m\right)+\frac{\frac{(1-\lambda)^3}{\lambda_\varepsilon^3}-k_m^3}{3}\right\} \\
&= \frac{1-\lambda}{4\left(1-\lambda-(1-\lambda)k_m\right)\left(1-\lambda+(1-\lambda)^2\right)}\left\{\frac{(1-\lambda)^3-(1-\lambda)^3k_m^3}{3\left(1-\lambda-(1-\lambda)k_m\right)}-(1-\lambda)^2k_m\right\} \\
&= \frac{1-\lambda}{4(1-\lambda)^2(1-k_m)(2-\lambda)}\left\{\frac{(1-\lambda)^3-(1-\lambda)^3k_m^3}{3(1-\lambda)(1-k_m)}-(1-\lambda)^2k_m\right\} \\
&= \frac{(1-\lambda)}{4(1-k_m)(2-\lambda)}\left\{\frac{1-k_m^3}{3(1-k_m)}-k_m\right\}
\end{aligned}$$

From the benchmark model $S_d^L(0)$ is obtained by evaluating dealers' surplus defined in (15) at $\varepsilon = 0$, yielding:

$$S_d^L(0) \quad = \quad \frac{(1-\lambda)^2}{12\left(1-\lambda+(1-\lambda)^2\right)} = \frac{(1-\lambda)}{12(2-\lambda)}$$

where $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ has been substituted out. Then $\gamma > S_d^L(0) - S_d(0)$ is:

$$\gamma \quad > \quad \frac{(1-\lambda)}{12(2-\lambda)} - \frac{(1-\lambda)}{4(1-k_m)(2-\lambda)}\left\{\frac{1-k_m^3}{3(1-k_m)}-k_m\right\}$$

$$
\begin{aligned}
&= \frac{(1-\lambda)}{4(2-\lambda)}\left\{\frac{1}{3}-\frac{1}{(1-k_m)}\left[\frac{1-k_m^3}{3(1-k_m)}-k_m\right]\right\} \\
&= \frac{(1-\lambda)}{4(2-\lambda)}\left\{\frac{1}{3}-\frac{1}{(1-k_m)}\left[\frac{1-k_m^3-3k_m+3k_m^2}{3(1-k_m)}\right]\right\} \\
&= \frac{(1-\lambda)}{4(2-\lambda)}\left\{\frac{1}{3}-\frac{(1-k_m)^3}{3(1-k_m)^2}\right\} = \frac{(1-\lambda)}{12(2-\lambda)}k_m
\end{aligned}
$$

A.4 Derivation of $\hat{k}$

The left hand side of (19) is larger than the right hand side of (20) only if

$$
\frac{1}{12\left(1-\lambda+\lambda_\varepsilon^2\right)}\left\{(1-\lambda)^2-\lambda_\varepsilon\left(1-\lambda-\lambda_\varepsilon k_m\right)\right\} \quad > \quad \frac{(1-\lambda)}{12(2-\lambda)}k_m
$$

Because $(1-\lambda+\lambda_\varepsilon^2)>0$, since $\lambda_\varepsilon = 1-\lambda+\varepsilon$ and $\varepsilon \in (0,\lambda)$, then this can be rearranged as

$$
\begin{aligned}
(2-\lambda)\left\{(1-\lambda)^2-\lambda_\varepsilon\left(1-\lambda-\lambda_\varepsilon k_m\right)\right\} &> (1-\lambda)\left(1-\lambda+\lambda_\varepsilon^2\right)k_m \\
(2-\lambda)(1-\lambda)(1-\lambda-\lambda_\varepsilon) &> \left[(1-\lambda)^2-\lambda_\varepsilon^2\right]k_m \\
(2-\lambda)(1-\lambda)(1-\lambda-\lambda_\varepsilon) &> (1-\lambda+\lambda_\varepsilon)(1-\lambda-\lambda_\varepsilon)k_m
\end{aligned}
$$

because $(1-\lambda-\lambda_\varepsilon)<0$ then we can rearranged the last inequality as

$$
(2-\lambda)(1-\lambda) \quad < \quad (1-\lambda+\lambda_\varepsilon)k_m
$$

which yields $k_m > \frac{(1-\lambda)(2-\lambda)}{(1-\lambda+\lambda_\varepsilon)} = \hat{k}$.

A.5 Derivation of $\hat{k} < \overline{k}$

From the definitions of $\hat{k} = \frac{(2-\lambda)}{2+\frac{\varepsilon}{(1-\lambda)}}$ and $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ we have that $\hat{k} < \overline{k}$ if and only if:

$$
(2-\lambda)\lambda_\varepsilon \quad < \quad 2(1-\lambda)+\varepsilon
$$

which, using the definition of $\lambda_\varepsilon = (1-\lambda+\varepsilon)$, can be rearranged as:

$$\begin{aligned} 2(1-\lambda) - \lambda(1-\lambda) &< 2(1-\lambda) + \varepsilon(\lambda-1) \\ -\lambda(1-\lambda) &< -\varepsilon(1-\lambda) \\ \lambda &> \varepsilon \end{aligned}$$

which is always true by the definition of $\varepsilon \in (0, \lambda)$.

A.6 Derivation of consumers' and producers' surplus with k_m

In order to obtain (22) consider the definition of consumers' surplus:

$$S_c(\underline{a}, \overline{a}; \varepsilon) = \frac{(1-\lambda)}{(\overline{a}-\underline{a})} \left[\int_{\underline{a}}^{\overline{a}} \int_{\underline{a}}^{v} (v-a)\, da dv + \int_{\overline{a}}^{1} \int_{\underline{a}}^{\overline{a}} (v-a)\, da dv \right]$$

for $\underline{a} = (k_m)$ and $\overline{a} = a\left(\overline{k}\right)$ with $a(k) = \frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$.

Consistently with the results in the previous sections we have $S_c\left(\overline{k}\right) = \frac{(1-\lambda)(1-a(0))^2}{6}$. This follows form $S_c(\underline{a}, \overline{a}; \varepsilon)$ evaluated at $\underline{a} = (0)$:

$$\begin{aligned} S_c(\underline{a}, \overline{a}; \varepsilon) &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})} \left[\int_{\underline{a}}^{\overline{a}} \left(v(v-\underline{a}) - \frac{(v^2-\underline{a}^2)}{2} \right) dv + \int_{\overline{a}}^{1} \left(v(\overline{a}-\underline{a}) - \frac{(\overline{a}^2-\underline{a}^2)}{2} \right) dv \right] \\ &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})} \left[\int_{\underline{a}}^{\overline{a}} \left(\frac{(v^2+\underline{a}^2)}{2} - v\underline{a} \right) dv + \int_{\overline{a}}^{1} \left(v(\overline{a}-\underline{a}) - \frac{(\overline{a}^2-\underline{a}^2)}{2} \right) dv \right] \\ &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})} \left[\frac{(\overline{a}^3-\underline{a}^3)}{6} - \underline{a}\frac{(\overline{a}^2-\underline{a}^2)}{2} + \frac{\underline{a}^2}{2}(\overline{a}-\underline{a}) + \frac{(1-\underline{a}^2)}{2}(\overline{a}-\underline{a}) - \frac{(\overline{a}^2-\underline{a}^2)}{2}(1-\overline{a}) \right] \\ &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})} \left[\frac{(\overline{a}^3-\underline{a}^3)}{6} - \frac{(\overline{a}^2-\underline{a}^2)}{2}(\underline{a}+1-\overline{a}) + \frac{(1-(\overline{a}^2-\underline{a}^2))}{2}(\overline{a}-\underline{a}) \right] \\ &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})} \left[\frac{(\overline{a}^3-\underline{a}^3)}{6} - \frac{(\overline{a}-\underline{a})(\overline{a}+\underline{a})}{2}(1-(\overline{a}-\underline{a})) + \frac{(1-(\overline{a}-\underline{a})(\overline{a}+\underline{a}))}{2}(\overline{a}-\underline{a}) \right] \\ &= (1-\lambda) \left[\frac{(\overline{a}^3-\underline{a}^3)}{6(\overline{a}-\underline{a})} + \frac{1-(\overline{a}-\underline{a})(\overline{a}+\underline{a}) - (\overline{a}+\underline{a})(1-(\overline{a}-\underline{a}))}{2} \right] \end{aligned}$$

$$
\begin{aligned}
&= (1-\lambda)\left[\frac{(\overline{a}^3-\underline{a}^3)}{6(\overline{a}-\underline{a})}+\frac{1-(\overline{a}+\underline{a})}{2}\right]\\
&= (1-\lambda)\left[\frac{(\overline{a}^2+\underline{a}^2+\overline{a}\underline{a})}{6}+\frac{1-(\overline{a}+\underline{a})}{2}\right]\\
&= \frac{(1-\lambda)}{6}\left[\overline{a}^2+\underline{a}^2+\overline{a}\underline{a}+3-3(\overline{a}+\underline{a})\right]\\
&= \frac{(1-\lambda)}{6}\left[\overline{a}(\overline{a}+\underline{a})+3-3(\overline{a}+\underline{a})+\underline{a}^2\right]\\
&= \frac{(1-\lambda)}{6}\left[3+(\overline{a}-3)(\overline{a}+\underline{a})+\underline{a}^2\right]
\end{aligned}
$$

Evaluating this at $\underline{a}=a(0)$ and $\overline{a}=a\left(\overline{k}\right)$ with $a(k)=\frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\overline{k}=\frac{1-\lambda}{\lambda_\varepsilon}$ yields:

$$
\begin{aligned}
\overline{a} &= 1\\
\underline{a} &= \frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}
\end{aligned}
$$

So that

$$
\begin{aligned}
S_c\left(a(0),a\left(\overline{k}\right);\varepsilon\right) &= \frac{(1-\lambda)}{6}\left[3-2\left(1+\frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right)+\left(\frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2\right]\\
&= \frac{(1-\lambda)}{6}\left[3-2\left(\frac{3(1-\lambda)+4\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right)+\frac{(1-\lambda+2\lambda_\varepsilon^2)^2}{4(1-\lambda+\lambda_\varepsilon^2)^2}\right]\\
&= \frac{(1-\lambda)}{12(1-\lambda+\lambda_\varepsilon^2)}\left[6\left(1-\lambda+\lambda_\varepsilon^2\right)-2\left(3(1-\lambda)+4\lambda_\varepsilon^2\right)+\frac{(1-\lambda+2\lambda_\varepsilon^2)^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right]\\
&= \frac{(1-\lambda)}{12(1-\lambda+\lambda_\varepsilon^2)}\left[\frac{(1-\lambda+2\lambda_\varepsilon^2)^2}{2(1-\lambda+\lambda_\varepsilon^2)}-2\lambda_\varepsilon^2\right]\\
&= \frac{(1-\lambda)}{12(1-\lambda+\lambda_\varepsilon^2)}\left[\frac{(1-\lambda)^2+4(1-\lambda)\lambda_\varepsilon^2+4\lambda_\varepsilon^4-4\lambda_\varepsilon^2(1-\lambda+\lambda_\varepsilon^2)}{2(1-\lambda+\lambda_\varepsilon^2)}\right]\\
&= \frac{(1-\lambda)^3}{24(1-\lambda+\lambda_\varepsilon^2)^2}
\end{aligned}
$$

This is the consumers' surplus from the low cost distribution, that is the same as in the benchmark model.

For the calculation of consumers' surplus with the high cost distribution we have instead:

$$S_c\left(a\left(k_m\right), a\left(\overline{k}\right); \varepsilon\right) = \frac{1-\lambda}{6}\left[3+\left(a\left(\overline{k}\right)+a\left(k_m\right)\right)\left(a\left(\overline{k}\right)-3\right)+a\left(k_m\right)^2\right]$$

using $a\left(\overline{k}\right)=1$ and letting $a\left(k_m\right)=\underline{a}$ as above, we then have

$$\begin{aligned} S_c\left(\underline{a}, 1; \varepsilon\right) &= \frac{1-\lambda}{6}\left[3-2\left(1+\underline{a}\right)+\underline{a}^2\right] \\ &= \frac{1-\lambda}{6}\left[1-2\underline{a}+\underline{a}^2\right]=\frac{(1-\lambda)}{6}\left(1-\underline{a}\right)^2 \end{aligned}$$

Using then $a(k)=\frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2\left(1-\lambda+\lambda_\varepsilon^2\right)}$ we have

$$\begin{aligned} S_c\left(\underline{a}, 1; \varepsilon\right) &= \frac{(1-\lambda)}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2\left(1-\lambda+\lambda_\varepsilon^2\right)-\left(1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon\right)\right]^2 \\ &= \frac{(1-\lambda)}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left(1-\lambda-k\lambda_\varepsilon\right)^2 \end{aligned}$$

from which it is easy to see that $S_c\left(\underline{a}, 1; \varepsilon\right)<S_c\left(a(0), 1; \varepsilon\right)=S_c^L(\varepsilon)$ for all $k_m>0$. In fact we have that the difference in consumers' surplus from investing in the low cost distribution is $S_c^L(\varepsilon)-S_c^H(\varepsilon)=S_c\left(a(0), 1; \varepsilon\right)-S_c\left(\underline{a}, 1; \varepsilon\right)$:

$$\begin{aligned} S_c^L(\varepsilon)-S_c^H(\varepsilon) &= \frac{(1-\lambda)^3}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}-\frac{(1-\lambda)}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left(1-\lambda-k_m\lambda_\varepsilon\right)^2 \\ &= \frac{(1-\lambda)}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[(1-\lambda)^2-\left(1-\lambda-k_m\lambda_\varepsilon\right)^2\right] \\ &= \frac{(1-\lambda)\,k_m\lambda_\varepsilon}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2(1-\lambda)-k_m\lambda_\varepsilon\right] \end{aligned}$$

which is always strictly positive because $2(1-\lambda)>k_m\lambda_\varepsilon$ since $\varepsilon\in(0,\lambda)$ and $k_m<\frac{1-\lambda}{\lambda_\varepsilon}$.

Similarly, for producers, we have that the bid price (9) is

$$b(k) = \lambda_\varepsilon\left(1-a(k)\right)=\lambda_\varepsilon\left(1-\frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2\left(1-\lambda+\lambda_\varepsilon^2\right)}\right)$$

$$= \quad \lambda_\varepsilon \frac{(1-\lambda) - k\lambda_\varepsilon}{2\left(1-\lambda+\lambda_\varepsilon^2\right)}$$

So that $\overline{b} = b(0) = \frac{\lambda_\varepsilon(1-\lambda)}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\underline{b} = b\left(\overline{k}\right) = 0$. Then, producers' surplus is

$$\begin{aligned}
S_p\left(\underline{b}, \overline{b}; \varepsilon\right) &= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\int_{\underline{b}}^{\overline{b}}\int_{v}^{\overline{b}}(b-v)\,dbdv + \int_{0}^{\underline{b}}\int_{\underline{b}}^{\overline{b}}(b-v)\,dbdv\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\int_{\underline{b}}^{\overline{b}}\frac{\overline{b}^2-v^2}{2} - v\left(\overline{b}-v\right)dv + \int_{0}^{\underline{b}}\frac{\overline{b}^2-\underline{b}^2}{2} - v\left(\overline{b}-\underline{b}\right)dv\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\frac{\overline{b}^2}{2}\left(\overline{b}-\underline{b}\right) + \frac{\left(\overline{b}^3-\underline{b}^3\right)}{6} - \overline{b}\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2} + \frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}\underline{b} - \frac{\underline{b}^2}{2}\left(\overline{b}-\underline{b}\right)\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}\left(\overline{b}-\underline{b}\right) + \frac{\left(\overline{b}^3-\underline{b}^3\right)}{6} - \left(\overline{b}-\underline{b}\right)\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\frac{\left(\overline{b}^3-\underline{b}^3\right)}{6} = \frac{\overline{b}^2+\overline{b}\underline{b}+\underline{b}^2}{6} = \frac{\overline{b}^2}{6}
\end{aligned}$$

Then, the gain in producers' surplus from dealers' investment into the low cost technology is

$$\begin{aligned}
S_p^L(0, b(0); \varepsilon) - S_p(0, b(k_m); \varepsilon) &= \frac{b(0)^2 - b(k_m)^2}{6} \\
&= \frac{\left(\lambda_\varepsilon\frac{(1-\lambda)}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2 - \left(\lambda_\varepsilon\frac{(1-\lambda-k_m\lambda_\varepsilon)}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2}{6} \\
&= \frac{\lambda_\varepsilon^2}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[(1-\lambda)^2 - (1-\lambda-k_m\lambda_\varepsilon)^2\right] \\
&= \frac{\lambda_\varepsilon^2}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2(1-\lambda)k_m\lambda_\varepsilon - (k_m\lambda_\varepsilon)^2\right] \\
&= \frac{\lambda_\varepsilon^3 k_m}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2(1-\lambda) - k_m\lambda_\varepsilon\right] > 0
\end{aligned}$$

where the last inequality follows from $k_m < \overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$. In fact

$$\begin{aligned} S_p^L(0, b(0); \varepsilon) - S_p(0, b(k_m); \varepsilon) &> S_p^L(0, b(0); \varepsilon) - S_p\left(0, b\left(\overline{k}\right); \varepsilon\right) \\ &= \frac{\lambda_\varepsilon^3 k_m [2(1-\lambda) - (1-\lambda)]}{24(1-\lambda+\lambda_\varepsilon^2)^2} > 0. \end{aligned}$$

B Model with ex-ante fixed investment

B.1 Derivation of $S_d(\varepsilon)$

$$\begin{aligned} S_d(\varepsilon) &= \int_{k_m}^{\overline{k}} \Pi(k; \lambda, \varepsilon) \frac{dk}{(1-k_m)} \\ &= \frac{1}{N(1-k_m)4(1-\lambda+\lambda_\varepsilon^2)} \int_{k_m}^{\overline{k}} (1-\lambda-k\lambda_\varepsilon)^2 dk \\ &= \frac{1}{\left(\overline{k}-k_m\right)(1-k_m)4(1-\lambda+\lambda_\varepsilon^2)} \int_{k_m}^{\overline{k}} (1-\lambda-k\lambda_\varepsilon)^2 dk \\ &= \frac{1}{\left(\overline{k}-k_m\right)(1-k_m)4(1-\lambda+\lambda_\varepsilon^2)} \int_{k_m}^{\overline{k}} \left((1-\lambda)^2 + (k\lambda_\varepsilon)^2 - 2(1-\lambda)k\lambda_\varepsilon\right) dk \\ &= \frac{1}{\left(\overline{k}-k_m\right)(1-k_m)4(1-\lambda+\lambda_\varepsilon^2)} \left\{(1-\lambda)^2\left(\overline{k}-k_m\right) - 2(1-\lambda)\lambda_\varepsilon \frac{\overline{k}^2-k_m^2}{2} + \lambda_\varepsilon^2 \frac{\overline{k}^3-k_m^3}{3}\right\} \\ &= \frac{1}{\left(\overline{k}-k_m\right)(1-k_m)4(1-\lambda+\lambda_\varepsilon^2)} \left\{(1-\lambda)^2 N - \lambda_\varepsilon(1-\lambda)\left(\overline{k}+k_m\right)N + \frac{\lambda_\varepsilon^2}{3}\left(\overline{k}^3-k_m^3\right)\right\} \quad (35) \end{aligned}$$

Thus

$$\begin{aligned} S_d(0) &= \frac{1}{N(1-k_m)4(1-\lambda)(2-\lambda)} \left\{(1-\lambda)^2 N - \lambda_\varepsilon(1-\lambda)\left(\overline{k}+k_m\right)N + \frac{(1-\lambda)^2}{3}\left(\overline{k}^3-k_m^3\right)\right\} \\ &= \frac{1-\lambda}{4N(1-k_m)(2-\lambda)} \left\{N\left(1-\overline{k}-k_m\right) + \frac{\left(\overline{k}^3-k_m^3\right)}{3}\right\} \quad (36) \end{aligned}$$

B.2 Derivation of (19)

Consider $S_d^L(\varepsilon) - S_d(\varepsilon) > \gamma$. Equation (15) can be rearranged as:

$$\begin{aligned}
S_d(\varepsilon) &= \frac{1}{(1-k_m)\,4\,(1-\lambda+\lambda_\varepsilon^2)}\left\{(1-\lambda)^2 - \lambda_\varepsilon(1-\lambda)\left(\overline{k}+k_m\right) + \frac{\lambda_\varepsilon^2}{3}\frac{\left(\overline{k}^3 - k_m^3\right)}{\left(\overline{k}-k_m\right)}\right\} \\
&= \frac{1}{(1-k_m)\,4\,(1-\lambda+\lambda_\varepsilon^2)}\left\{(1-\lambda)^2 - \lambda_\varepsilon(1-\lambda)\left(\overline{k}+k_m\right) + \frac{\lambda_\varepsilon^2}{3}\left(\overline{k}^2 + \overline{k}k_m + k_m^2\right)\right\}
\end{aligned}$$

Substituting out $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ yields

$$\begin{aligned}
S_d(\varepsilon) &= \frac{\left\{(1-\lambda)^2 - \lambda_\varepsilon(1-\lambda)\left(\frac{(1-\lambda)}{\lambda_\varepsilon}+k_m\right) + \frac{\lambda_\varepsilon^2}{3}\left(\frac{(1-\lambda)^2}{\lambda_\varepsilon^2} + \frac{(1-\lambda)}{\lambda_\varepsilon}k_m + k_m^2\right)\right\}}{(1-k_m)\,4\,(1-\lambda+\lambda_\varepsilon^2)} \\
&= \frac{\left\{(1-\lambda)^2 - (1-\lambda)^2 - \lambda_\varepsilon(1-\lambda)k_m + \frac{(1-\lambda)^2}{3} + \frac{\lambda_\varepsilon}{3}(1-\lambda)k_m + \frac{\lambda_\varepsilon^2}{3}k_m^2\right\}}{(1-k_m)\,4\,(1-\lambda+\lambda_\varepsilon^2)} \\
&= \frac{1}{(1-k_m)\,4\,(1-\lambda+\lambda_\varepsilon^2)}\left\{-\frac{2\lambda_\varepsilon}{3}(1-\lambda)k_m + \frac{(1-\lambda)^2}{3} + \frac{\lambda_\varepsilon^2}{3}k_m^2\right\} \\
&= \frac{1}{12\,(1-k_m)(1-\lambda+\lambda_\varepsilon^2)}\left\{-2\lambda_\varepsilon(1-\lambda)k_m + (1-\lambda)^2 + \lambda_\varepsilon^2 k_m^2\right\} \\
&= \frac{1}{12\,(1-k_m)(1-\lambda+\lambda_\varepsilon^2)}\left\{\lambda_\varepsilon k_m(\lambda_\varepsilon k_m - 2(1-\lambda)) + (1-\lambda)^2\right\} \qquad (37)
\end{aligned}$$

Thus $S_d^L(\varepsilon) - S_d^H(\varepsilon) > \gamma$ if and only if

$$\frac{(1-\lambda)^2}{12\,(1-\lambda+\lambda_\varepsilon^2)} - \frac{\left\{\lambda_\varepsilon k_m(\lambda_\varepsilon k_m - 2(1-\lambda)) + (1-\lambda)^2\right\}}{12\,(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} > \gamma$$

which can be rewritten as

$$\begin{aligned}
\frac{(1-k_m)(1-\lambda)^2 - \lambda_\varepsilon k_m(\lambda_\varepsilon k_m - 2(1-\lambda)) - (1-\lambda)^2}{12\,(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} &> \gamma \\
k_m\frac{-(1-\lambda)^2 - \lambda_\varepsilon(\lambda_\varepsilon k_m - 2(1-\lambda))}{12\,(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} &> \gamma
\end{aligned}$$

$$k_m \frac{2\lambda_\varepsilon (1-\lambda) - (1-\lambda)^2 - \lambda_\varepsilon^2 k_m}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} \quad > \quad \gamma$$

which is (19).

Just to ease interpretation with respect to the lower extreme on the support of the high cost technology for dealers, k_m, we can express (19) as a sufficient condition on k_m as a function of γ. To do so, rearrange (19) as

$$-\lambda_\varepsilon^2 k_m^2 + \left[12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma + (1-\lambda)(2\lambda_\varepsilon - (1-\lambda))\right] k_m - 12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma \quad > \quad 0$$

which is violated for $k_m = 0$, and for $k_m = 1$ it becomes

$$2 \quad > \quad \lambda_\varepsilon (1-\lambda) = (1-\lambda)^2 + \varepsilon (1-\lambda)$$

the largest value that the right hand side can take is

$$(1-\lambda)^2 + \lambda(1-\lambda) \quad = \quad (1-\lambda)$$

thus the inequality is always satisfied at $k_m = 1$. For $k_m = \overline{k}$ it becomes

$$\begin{aligned} \left[12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma - (1-\lambda)^2 + (1-\lambda)\,2\lambda_\varepsilon\right]\frac{(1-\lambda)}{\lambda_\varepsilon} - 12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma - (1-\lambda)^2 &\quad > \quad 0 \\ \left[12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma - (1-\lambda)^2\right]\left(\frac{(1-\lambda)}{\lambda_\varepsilon} - 1\right) &\quad > \quad 0 \end{aligned}$$

which, because $\frac{(1-\lambda)}{\lambda_\varepsilon} < 1$, is satisfied if and only if

$$\gamma \quad < \quad \frac{(1-\lambda)^2}{12(1-\lambda+\lambda_\varepsilon^2)}$$

Then $S_d^L(\varepsilon) - S_d(\varepsilon) > \gamma$ for all $k_m \in (k_1, k_2)$ with

$$k_1(\gamma) \quad = \quad \frac{-\left[12(1-\lambda+\lambda_\varepsilon^2)\gamma + (1-\lambda)(2\lambda_\varepsilon - (1-\lambda))\right]}{-2\lambda_\varepsilon^2} +$$

$$
\begin{aligned}
k_2(\gamma) &= -\frac{\sqrt{\left[12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma+(1-\lambda)\left(2\lambda_\varepsilon-(1-\lambda)\right)\right]^2-4\lambda_\varepsilon^2 12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma}}{-2\lambda_\varepsilon^2} \\
&= \frac{-\left[12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma+(1-\lambda)\left(2\lambda_\varepsilon-(1-\lambda)\right)\right]}{-2\lambda_\varepsilon^2}+ \\
&\quad +\frac{\sqrt{\left[12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma+(1-\lambda)\left(2\lambda_\varepsilon-(1-\lambda)\right)\right]^2-4\lambda_\varepsilon^2 12\left(1-\lambda+\lambda_\varepsilon^2\right)\gamma}}{-2\lambda_\varepsilon^2}
\end{aligned}
$$

B.3 Derivation of (20)

Consider $\gamma > S_d^L(0) - S_d(0)$. Equation (15) can be rearranged using $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ as (37), which, evaluated at $\varepsilon = 0$ yields:

$$
\begin{aligned}
S_d(0) &= \frac{1}{12\left(1-k_m\right)\left(1-\lambda+(1-\lambda)^2\right)}\left\{(1-\lambda)k_m\left((1-\lambda)k_m-2(1-\lambda)\right)+(1-\lambda)^2\right\} \\
&= \frac{(1-\lambda)\left(1-2k_m+k_m^2\right)}{12\left(1-k_m\right)(2-\lambda)} = \frac{(1-\lambda)\left(1-k_m\right)}{12(2-\lambda)}
\end{aligned}
$$

From the benchmark model $S_d^L(0)$ is obtained by evaluating dealers' surplus defined in (37) at $k_m = 0$ and $\varepsilon = 0$, yielding:

$$
S_d^L(0) = \frac{(1-\lambda)}{12(2-\lambda)}
$$

Then $\gamma > S_d^L(0) - S_d(0)$ is:

$$
\begin{aligned}
\gamma &> \frac{(1-\lambda)}{12(2-\lambda)} - \frac{(1-\lambda)\left(1-k_m\right)}{12(2-\lambda)} \\
&= \frac{(1-\lambda)}{12(2-\lambda)}k_m
\end{aligned}
$$

which is (20).

B.4 Derivation of $\hat{k}$

The left hand side of (19) is larger than the right hand side of (20) only if

$$\frac{(1-\lambda)(2\lambda_\varepsilon-(1-\lambda))k_m-(\lambda_\varepsilon k_m)^2}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} > \frac{(1-\lambda)}{12(2-\lambda)}k_m \tag{38}$$

Because $(1-\lambda+\lambda_\varepsilon^2)>0$, since $\lambda_\varepsilon = 1-\lambda+\varepsilon$ and $\varepsilon \in (0,\lambda)$, then this can be rearranged as

$$\left[(1-\lambda)(2\lambda_\varepsilon-(1-\lambda))k_m-\lambda_\varepsilon^2 k_m^2\right] > \frac{(1-\lambda)}{(2-\lambda)}\left(1-\lambda+\lambda_\varepsilon^2\right)\left(k_m-k_m^2\right)$$

that can be rewritten as

$$\begin{aligned}\left[(1-\lambda)(2\lambda_\varepsilon-(1-\lambda))-\frac{(1-\lambda)^2}{(2-\lambda)}-\lambda_\varepsilon^2\frac{(1-\lambda)}{(2-\lambda)}\right]k_m+\left(\frac{(1-\lambda)^2}{(2-\lambda)}+\lambda_\varepsilon^2\left(\frac{(1-\lambda)}{(2-\lambda)}-1\right)\right)k_m^2 &> 0\\ \left[2\lambda_\varepsilon-\frac{(3-\lambda)(1-\lambda)+\lambda_\varepsilon^2}{(2-\lambda)}\right](1-\lambda)k_m+\left(\frac{(1-\lambda)^2-\lambda_\varepsilon^2}{(2-\lambda)}\right)k_m^2 &> 0\end{aligned} \tag{39}$$

Lemma 7. *The first term in square brackets in (39),* $\left[2\lambda_\varepsilon-\frac{(3-\lambda)(1-\lambda)+\lambda_\varepsilon^2}{(2-\lambda)}\right]$*, is always positive.*

Proof. Rearrange $\left[2\lambda_\varepsilon-\frac{(3-\lambda)(1-\lambda)+\lambda_\varepsilon^2}{(2-\lambda)}\right]$ as

$$-\lambda_\varepsilon^2+2(2-\lambda)\lambda_\varepsilon-(3-\lambda)(1-\lambda) > 0$$

which is satisfied for all $\lambda_\varepsilon \in (x_1,x_2)$ where

$$\begin{aligned}x_2 &= \frac{-(2-\lambda)-\sqrt{(2-\lambda)^2-(3-\lambda)(1-\lambda)}}{-1}\\ &= (2-\lambda)+\sqrt{(4-4\lambda+\lambda^2)-3-4\lambda+\lambda^2}\\ &= (2-\lambda)+1=(3-\lambda)>1\\ x_1 &= (2-\lambda)-1=(1-\lambda)\end{aligned}$$

Because $\lambda_\varepsilon \in ((1-\lambda), 1)$, and because $x_2 > 1$, then it is always the case that $\lambda_\varepsilon \in (x_1, x_2)$. □

We use this result in the following argument to characterize the values of k_m such that (39) is satisfied.

Let $f(k) = \left[2\lambda_\varepsilon - \frac{(3-\lambda)(1-\lambda)+\lambda_\varepsilon^2}{(2-\lambda)}\right](1-\lambda)k + \left(\frac{(1-\lambda)^2-\lambda_\varepsilon^2}{(2-\lambda)}\right)k^2$, that is the left hand side of (39) as a function of k. Then $f(k) = 0$ for $k = k_1 = 0$ and for $k = k_2$ defined by

$$\left[\frac{2(2-\lambda)\lambda_\varepsilon - (3-\lambda)(1-\lambda) - \lambda_\varepsilon^2}{(2-\lambda)}\right](1-\lambda) = \frac{\lambda_\varepsilon^2 - (1-\lambda)^2}{(2-\lambda)}k_2$$

which can be rewritten as

$$\left[\frac{2(2-\lambda)\lambda_\varepsilon - (3-\lambda)(1-\lambda) - \lambda_\varepsilon^2}{\lambda_\varepsilon^2 - (1-\lambda)^2}\right](1-\lambda) = k_2 \tag{40}$$

Notice that $k_2 > 0$ because $\lambda_\varepsilon > (1-\lambda)$ and because, by lemma 7, the numerator in the definition of k_2 is strictly positive.

Thus, (38) is satisfied for any $k_m \in (k_1, k_2)$, with $k_1 = 0$ and k_2 defined in (40). Because by definition $k_m > 0$ then the only relevant constraint on k_m is $k_m < k_2$. Let $\hat{k} = k_2$ defined in (40) and we have the result.

B.5 Properties of $\hat{k}$

Lemma 8. *Let $\hat{k}$ be defined in (17) and $\overline{k}_\varepsilon = \frac{1-\lambda}{\lambda_\varepsilon}$, $\overline{k}_0 = 1$. Then $k_m \in (\overline{k}_\varepsilon, \overline{k}_0)$ for all $\lambda \in (0,1)$ and $\varepsilon \in (0, \lambda)$.*

Proof. Consider economies with $\varepsilon > 0$. In this case $\overline{k} = \overline{k}_\varepsilon = \frac{1-\lambda}{\lambda_\varepsilon}$. From (17) and the definition of $\overline{k}$, it follows that $\hat{k} > \overline{k}$ if and only if

$$\left[\frac{2(2-\lambda)\lambda_\varepsilon - (3-\lambda)(1-\lambda) - \lambda_\varepsilon^2}{\lambda_\varepsilon^2 - (1-\lambda)^2}\right](1-\lambda) > \frac{(1-\lambda)}{\lambda_\varepsilon}$$

which, because $\lambda_\varepsilon^2 - (1-\lambda)^2 > 0$, can be rearranged as

$$\lambda_\varepsilon\left[2(2-\lambda)\lambda_\varepsilon - (3-\lambda)(1-\lambda) - \lambda_\varepsilon^2\right] > \lambda_\varepsilon^2 - (1-\lambda)^2$$

and further as

$$\left[-\left(\lambda_\varepsilon-(1-\lambda)\right)^2\right](\lambda_\varepsilon-1) \quad > \quad 0$$

Because by definition of λ_ε we have $\lambda_\varepsilon \in ((1-\lambda),1)$ then the above inequality is always satisfied. Consider now economies with $\varepsilon=0$. In this case $\overline{k}=\overline{k}_0=1$. Thus $\hat{k}>1$ if and only if

$$\left[2(2-\lambda)\lambda_\varepsilon-(3-\lambda)(1-\lambda)-\lambda_\varepsilon^2\right](1-\lambda) \quad > \quad \lambda_\varepsilon^2-(1-\lambda)^2$$

which, substituting out $\lambda_\varepsilon^2 1-\lambda+\varepsilon$, simplifies to

$$(1-\lambda)^2+2\varepsilon(1-\lambda) \quad > \quad (1-\lambda)^2+2\varepsilon(1-\lambda)+\varepsilon^2$$

The above inequality is never satisfied. □

B.6 Derivation of consumers' and producers' surplus with k_m

B.6.1 Consumers: low cost technology

In order to obtain (22) consider the definition of consumers' surplus:

$$S_c(\underline{a},\overline{a};\varepsilon) \quad = \quad \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\int_{\underline{a}}^{\overline{a}}\int_{\underline{a}}^{v}(v-a)\,dadv+\int_{\overline{a}}^{1}\int_{\underline{a}}^{\overline{a}}(v-a)\,dadv\right]$$

for $\underline{a}=a(k_m)$ and $\overline{a}=a\left(\overline{k}\right)$ with $a(k)=\frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$.

Consistently with the results in the previous sections we have $S_c\left(\overline{k}\right)=\frac{(1-\lambda)(1-a(0))^2}{6}$. This follows form $S_c(\underline{a},\overline{a};\varepsilon)$ evaluated at $\underline{a}=(0)$:

$$\begin{aligned} S_c(\underline{a},\overline{a};\varepsilon) &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\int_{\underline{a}}^{\overline{a}}\left(v(v-\underline{a})-\frac{(v^2-\underline{a}^2)}{2}\right)dv+\int_{\overline{a}}^{1}\left(v(\overline{a}-\underline{a})-\frac{(\overline{a}^2-\underline{a}^2)}{2}\right)dv\right] \\ &= \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\int_{\underline{a}}^{\overline{a}}\left(\frac{(v^2+\underline{a}^2)}{2}-v\underline{a}\right)dv+\int_{\overline{a}}^{1}\left(v(\overline{a}-\underline{a})-\frac{(\overline{a}^2-\underline{a}^2)}{2}\right)dv\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\frac{(\overline{a}^3-\underline{a}^3)}{6} - \underline{a}\frac{(\overline{a}^2-\underline{a}^2)}{2} + \frac{\underline{a}^2}{2}(\overline{a}-\underline{a}) + \frac{(1-\underline{a}^2)}{2}(\overline{a}-\underline{a}) - \frac{(\overline{a}^2-\underline{a}^2)}{2}(1-\overline{a})\right] \\
&= \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\frac{(\overline{a}^3-\underline{a}^3)}{6} - \frac{(\overline{a}^2-\underline{a}^2)}{2}(\underline{a}+1-\overline{a}) + \frac{(1-(\overline{a}^2-\underline{a}^2))}{2}(\overline{a}-\underline{a})\right] \\
&= \frac{(1-\lambda)}{(\overline{a}-\underline{a})}\left[\frac{(\overline{a}^3-\underline{a}^3)}{6} - \frac{(\overline{a}-\underline{a})(\overline{a}+\underline{a})}{2}(1-(\overline{a}-\underline{a})) + \frac{(1-(\overline{a}-\underline{a})(\overline{a}+\underline{a}))}{2}(\overline{a}-\underline{a})\right] \\
&= (1-\lambda)\left[\frac{(\overline{a}^3-\underline{a}^3)}{6(\overline{a}-\underline{a})} + \frac{1-(\overline{a}-\underline{a})(\overline{a}+\underline{a})-(\overline{a}+\underline{a})(1-(\overline{a}-\underline{a}))}{2}\right] \\
&= (1-\lambda)\left[\frac{(\overline{a}^3-\underline{a}^3)}{6(\overline{a}-\underline{a})} + \frac{1-(\overline{a}+\underline{a})}{2}\right] \\
&= (1-\lambda)\left[\frac{(\overline{a}^2+\underline{a}^2+\overline{a}\underline{a})}{6} + \frac{1-(\overline{a}+\underline{a})}{2}\right] \\
&= \frac{(1-\lambda)}{6}\left[\overline{a}^2+\underline{a}^2+\overline{a}\underline{a}+3-3(\overline{a}+\underline{a})\right] \\
&= \frac{(1-\lambda)}{6}\left[\overline{a}(\overline{a}+\underline{a})+3-3(\overline{a}+\underline{a})+\underline{a}^2\right] \\
&= \frac{(1-\lambda)}{6}\left[3+(\overline{a}-3)(\overline{a}+\underline{a})+\underline{a}^2\right]
\end{aligned}
$$

Evaluating this at $\underline{a} = a(0)$ and $\overline{a} = a\left(\overline{k}\right)$ with $a(k) = \frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ yields:

$$
\begin{aligned}
\overline{a} &= 1 \\
\underline{a} &= \frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}
\end{aligned}
$$

So that

$$
\begin{aligned}
S_c\left(a(0), a\left(\overline{k}\right); \varepsilon\right) &= \frac{(1-\lambda)}{6}\left[3-2\left(1+\frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right)+\left(\frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2\right] \\
&= \frac{(1-\lambda)}{6}\left[3-2\left(\frac{3(1-\lambda)+4\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right)+\frac{(1-\lambda+2\lambda_\varepsilon^2)^2}{4(1-\lambda+\lambda_\varepsilon^2)^2}\right] \\
&= \frac{(1-\lambda)}{12(1-\lambda+\lambda_\varepsilon^2)}\left[6\left(1-\lambda+\lambda_\varepsilon^2\right)-2\left(3(1-\lambda)+4\lambda_\varepsilon^2\right)+\frac{(1-\lambda+2\lambda_\varepsilon^2)^2}{2(1-\lambda+\lambda_\varepsilon^2)}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(1-\lambda)}{12\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)}\left[\frac{\left(1-\lambda+2\lambda_{\varepsilon}^{2}\right)^{2}}{2\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)}-2\lambda_{\varepsilon}^{2}\right] \\
&= \frac{(1-\lambda)}{12\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)}\left[\frac{(1-\lambda)^{2}+4(1-\lambda)\lambda_{\varepsilon}^{2}+4\lambda_{\varepsilon}^{4}-4\lambda_{\varepsilon}^{2}\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)}{2\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)}\right] \\
&= \frac{(1-\lambda)^{3}}{24\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)^{2}}
\end{aligned}
$$

This is the consumers' surplus from the low cost distribution, that is the same as in the benchmark model.

B.6.2 Consumers: high cost technology

For the calculation of consumers' surplus with the high cost technology we have instead:

$$
S_c\left(a\left(k_m\right), a\left(\overline{k}\right); \varepsilon\right) = \frac{1-\lambda}{6}\left[3+\left(a\left(\overline{k}\right)+a\left(k_m\right)\right)\left(a\left(\overline{k}\right)-3\right)+a\left(k_m\right)^{2}\right]
$$

using $a\left(\overline{k}\right)=1$ and letting $a\left(k_m\right)=\underline{a}$ as above, we then have

$$
\begin{aligned}
S_c\left(\underline{a}, 1; \varepsilon\right) &= \frac{1-\lambda}{6}\left[3-2\left(1+\underline{a}\right)+\underline{a}^{2}\right] \\
&= \frac{1-\lambda}{6}\left[1-2\underline{a}+\underline{a}^{2}\right] = \frac{(1-\lambda)}{6}\left(1-\underline{a}\right)^{2}
\end{aligned}
$$

Using then $a(k)=\frac{1-\lambda+2\lambda_{\varepsilon}^{2}+k\lambda_{\varepsilon}}{2\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)}$ we have

$$
\begin{aligned}
S_c\left(\underline{a}, 1; \varepsilon\right) &= \frac{(1-\lambda)}{24\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)^{2}}\left[2\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)-\left(1-\lambda+2\lambda_{\varepsilon}^{2}+k\lambda_{\varepsilon}\right)\right]^{2} \\
&= \frac{(1-\lambda)}{24\left(1-\lambda+\lambda_{\varepsilon}^{2}\right)^{2}}\left(1-\lambda-k\lambda_{\varepsilon}\right)^{2}
\end{aligned}
$$

from which it is easy to see that $S_c\left(\underline{a}, 1; \varepsilon\right) < S_c\left(a(0), 1; \varepsilon\right)$ for all $k_m > 0$, where, we denote the consumers' surplus for a given ε in the economy with the low and high cost technologies, respectively, as $S_c^L(\varepsilon) = S_c\left(a(0), 1; \varepsilon\right)$ and $S_c^H(\varepsilon) = S_c\left(\underline{a}, 1; \varepsilon\right)$. In fact, the difference in consumers' surplus from investing in the low cost technology is $S_c^L(\varepsilon) - S_c^H(\varepsilon) =$

$S_c\left(a\left(0\right),1;\varepsilon\right)-S_c\left(\underline{a},1;\varepsilon\right)$:

$$\begin{aligned}
S_c^L\left(\varepsilon\right)-S_c^H\left(\varepsilon\right) &= \frac{\left(1-\lambda\right)^3}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}-\frac{\left(1-\lambda\right)}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left(1-\lambda-k_m\lambda_\varepsilon\right)^2 \\
&= \frac{\left(1-\lambda\right)}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[\left(1-\lambda\right)^2-\left(1-\lambda-k_m\lambda_\varepsilon\right)^2\right] \\
&= \frac{\left(1-\lambda\right)k_m\lambda_\varepsilon}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2\left(1-\lambda\right)-k_m\lambda_\varepsilon\right]
\end{aligned}$$

which is always strictly positive because $2\left(1-\lambda\right)>k_m\lambda_\varepsilon$ since $\varepsilon\in\left(0,\lambda\right)$ and $k_m<\frac{1-\lambda}{\lambda_\varepsilon}$.

B.6.3 Producers

Similarly, for producers, we have that the bid price is

$$\begin{aligned}
b\left(k\right) &= \lambda_\varepsilon\left(1-a\left(k\right)\right)=\lambda_\varepsilon\left(1-\frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2\left(1-\lambda+\lambda_\varepsilon^2\right)}\right) \\
&= \lambda_\varepsilon\frac{\left(1-\lambda\right)-k\lambda_\varepsilon}{2\left(1-\lambda+\lambda_\varepsilon^2\right)}
\end{aligned}$$

So that $\overline{b}=b\left(0\right)=\frac{\lambda_\varepsilon(1-\lambda)}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\underline{b}=b\left(\overline{k}\right)=0$. Then, producers' surplus is

$$\begin{aligned}
S_p\left(\underline{b},\overline{b};\varepsilon\right) &= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\int_{\underline{b}}^{\overline{b}}\int_v^{\overline{b}}\left(b-v\right)dbdv+\int_0^{\underline{b}}\int_{\underline{b}}^{\overline{b}}\left(b-v\right)dbdv\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\int_{\underline{b}}^{\overline{b}}\frac{\overline{b}^2-v^2}{2}-v\left(\overline{b}-v\right)dv+\int_0^{\underline{b}}\frac{\overline{b}^2-\underline{b}^2}{2}-v\left(\overline{b}-\underline{b}\right)dv\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\frac{\overline{b}^2}{2}\left(\overline{b}-\underline{b}\right)+\frac{\left(\overline{b}^3-\underline{b}^3\right)}{6}-\overline{b}\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}+\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}\underline{b}-\frac{\underline{b}^2}{2}\left(\overline{b}-\underline{b}\right)\right] \\
&= \frac{1}{\left(\overline{b}-\underline{b}\right)}\left[\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}\left(\overline{b}-\underline{b}\right)+\frac{\left(\overline{b}^3-\underline{b}^3\right)}{6}-\left(\overline{b}-\underline{b}\right)\frac{\left(\overline{b}^2-\underline{b}^2\right)}{2}\right]
\end{aligned}$$

$$= \quad \frac{1}{\left(\overline{b}-\underline{b}\right)}\frac{\left(\overline{b}^3-\underline{b}^3\right)}{6}=\frac{\overline{b}^2+\overline{b}\underline{b}+\underline{b}^2}{6}=\frac{\overline{b}^2}{6}$$

Let $S_p^L(\varepsilon) = S_p\left(0, b\left(0\right); \varepsilon\right)$ and $S_p^H(\varepsilon) = S_p\left(0, b\left(k_m\right); \varepsilon\right)$ denote the producers' surplus, for a given ε, in the economy with the low and high cost technologies respectively. Then, the gain in producers' surplus from dealers' investment into the low cost technology is simply:

$$\begin{aligned}
S_p^L\left(0, b\left(0\right); \varepsilon\right) - S_p\left(0, b\left(k_m\right); \varepsilon\right) &= \frac{b\left(0\right)^2 - b\left(k_m\right)^2}{6} \\
&= \frac{\left(\lambda_\varepsilon \frac{(1-\lambda)}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2 - \left(\lambda_\varepsilon \frac{(1-\lambda-k_m\lambda_\varepsilon)}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2}{6} \\
&= \frac{\lambda_\varepsilon^2}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[\left(1-\lambda\right)^2 - \left(1-\lambda-k_m\lambda_\varepsilon\right)^2\right] \\
&= \frac{\lambda_\varepsilon^2}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2\left(1-\lambda\right)k_m\lambda_\varepsilon - \left(k_m\lambda_\varepsilon\right)^2\right] \\
&= \frac{\lambda_\varepsilon^3 k_m}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2\left(1-\lambda\right) - k_m\lambda_\varepsilon\right] > 0
\end{aligned}$$

where the last inequality follows from $k_m < \overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$. In fact

$$\begin{aligned}
S_p^L\left(0, b\left(0\right); \varepsilon\right) - S_p\left(0, b\left(k_m\right); \varepsilon\right) &> S_p^L\left(0, b\left(0\right); \varepsilon\right) - S_p\left(0, b\left(\overline{k}\right); \varepsilon\right) \\
&= \frac{\lambda_\varepsilon^3 k_m\left[2\left(1-\lambda\right) - \left(1-\lambda\right)\right]}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2} > 0.
\end{aligned}$$

B.7 Derivation of social planner's investment choice in (30)

In an economy with idiosyncratic risk (i.e. $\varepsilon > 0$) the social planner invests if and only if:

$$S_c^L\left(\varepsilon\right) - S_c^H\left(\varepsilon\right) + S_p^L\left(\varepsilon\right) - S_p^H\left(\varepsilon\right) + S_d^L\left(\varepsilon\right) - S_d^H\left(\varepsilon\right) \quad > \quad \gamma$$

That is to say

$$\frac{\left(1-\lambda\right)k_m\lambda_\varepsilon}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2\left(1-\lambda\right) - k_m\lambda_\varepsilon\right] + \frac{\lambda_\varepsilon^3 k_m}{24\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\left[2\left(1-\lambda\right) - k_m\lambda_\varepsilon\right] \quad +$$

$$\frac{(1-\lambda)(2\lambda_\varepsilon-(1-\lambda))k_m-(\lambda_\varepsilon k_m)^2}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} > \gamma$$

which can be rearranged as

$$\begin{aligned}\frac{k_m\lambda_\varepsilon[2(1-\lambda)-k_m\lambda_\varepsilon][(1-\lambda)+\lambda_\varepsilon^2]}{24(1-\lambda+\lambda_\varepsilon^2)^2} & + \\ \frac{(1-\lambda)k_m 2\lambda_\varepsilon-(1-\lambda)^2k_m-(\lambda_\varepsilon k_m)^2}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} & > \gamma\end{aligned}$$

and further as

$$\begin{aligned}\frac{k_m\lambda_\varepsilon[2(1-\lambda)-k_m\lambda_\varepsilon]}{24(1-\lambda+\lambda_\varepsilon^2)} & + \\ \frac{(1-\lambda)k_m 2\lambda_\varepsilon-(1-\lambda)^2k_m-(\lambda_\varepsilon k_m)^2}{12(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} & > \gamma\end{aligned}$$

and

$$\begin{aligned}\frac{(1-k_m)k_m\lambda_\varepsilon[2(1-\lambda)-k_m\lambda_\varepsilon]}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} & + \\ \frac{(1-\lambda)k_m 4\lambda_\varepsilon-2(1-\lambda)^2k_m-2(\lambda_\varepsilon k_m)^2}{24(1-k_m)(1-\lambda+\lambda_\varepsilon^2)} & > \gamma\end{aligned}$$

C Average bid-ask spreads

Proof of Proposition 6.

Proof. The average bid ask spread is:

$$\begin{aligned}s^L(0,\varepsilon)-s^H(k_m,0) = & \int_{\underline{a}^L}^{\overline{a}^L} a\frac{da}{\overline{a}^L-\underline{a}^L}-\int_{\underline{a}^H}^{\overline{a}^H} a\frac{da}{\overline{a}^H-\underline{a}^H}- \\ & \left[\int_{\underline{b}^L}^{\overline{b}^L} b\frac{db}{\overline{b}^L-\underline{b}^L}-\int_{\underline{b}^H}^{\overline{b}^H} b\frac{db}{\overline{b}^H-\underline{b}^H}\right]\end{aligned}$$

where $\overline{a}^L=\overline{a}^H=1$ and $\underline{b}^L=\underline{b}^H=0$ because these are the prices that the least efficient dealer charges, which is dealer $k = \overline{k}$ in the economy without insurance and the low cost technology, and it is dealer $k = 1$ in the economy with insurance and the high cost technology. Then the average bid ask spread is

$$
\begin{aligned}
s^L(0,\varepsilon) - s^H(k_m, 0) &= \int_{\underline{a}^L}^{1} a \frac{da}{1-\underline{a}^L} - \int_{\underline{a}^H}^{1} a \mid_{\varepsilon=0} \frac{da}{1-\underline{a}^H} - \\
&\quad \left[\int_0^{\bar{b}^L} b \frac{db}{\bar{b}^L} - \int_0^{\bar{b}^H} b \mid_{\varepsilon=0} \frac{db}{\bar{b}^H} \right]
\end{aligned}
$$

where $\underline{a} = a(k_m)$, with $a(k) = \frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, implies that

$$
\begin{aligned}
\underline{a}^H = a(k_m, \varepsilon = 0) &= \frac{(1-\lambda) + 2(1-\lambda)^2 + k_m(1-\lambda)}{2(1-\lambda)(2-\lambda)} \\
&= \frac{1 + 2(1-\lambda) + k_m}{2(2-\lambda)} \\
&= \frac{(3-2\lambda) + k_m}{2(2-\lambda)}
\end{aligned}
$$

where we also used the fact that $\varepsilon = 0$ because there is insurance in the economy with the high cost technology.

Similarly $\underline{a}^L = a(0, \varepsilon > 0)$, which, with with $a(k) = \frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\varepsilon > 0$ because there is no insurance in the economy with the low cost technology, implies that

$$
\begin{aligned}
\underline{a}^L &= \frac{1-\lambda+2\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)} \\
&= \frac{1}{2} + \frac{\lambda_\varepsilon^2}{2(1-\lambda+\lambda_\varepsilon^2)}
\end{aligned}
$$

Analogously the bid price can be rearranged as: $b(k) = \lambda_\varepsilon(1 - a(k)) = \lambda_\varepsilon \left(1 - \frac{1-\lambda+2\lambda_\varepsilon^2+k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}\right)$ which, for the economy with the high cost technology and insurance, implies $\bar{b}^H = (1-\lambda)(1 - a(k_m, \varepsilon = 0))$ that yields

$$
\bar{b}^H = (1-\lambda)\left(1 - \underline{a}^H\right) = \frac{(1-\lambda)(1-k_m)}{2(2-\lambda)}
$$

and

$$\begin{aligned}
\overline{b}^L &= \lambda_\varepsilon \left(1 - a\left(0, \varepsilon > 0\right)\right) = \lambda_\varepsilon \left(1 - \underline{a}^L\right) \\
&= \lambda_\varepsilon \left(1 - \frac{1 - \lambda + 2\lambda_\varepsilon^2}{2\left(1 - \lambda + \lambda_\varepsilon^2\right)}\right) \\
&= \lambda_\varepsilon \frac{(1 - \lambda)}{2\left(1 - \lambda + \lambda_\varepsilon^2\right)}
\end{aligned}$$

Therefore, we have that the difference in average bid ask spreads is

$$\begin{aligned}
s^L(0, \varepsilon) - s^H(k_m, 0) &= \int_{\underline{a}^L}^1 a \frac{da}{1 - \underline{a}^L} - \int_{\underline{a}^H}^1 a \mid_{\varepsilon=0} \frac{da}{1 - \underline{a}^H} - \left[\int_0^{\overline{b}^L} b \frac{db}{\overline{b}^L} - \int_0^{\overline{b}^H} b \mid_{\varepsilon=0} \frac{db}{\overline{b}^H}\right] \\
&= \frac{1 - \left(\underline{a}^L\right)^2}{2\left(1 - \underline{a}^L\right)} - \left(\frac{1 - \left(\underline{a}^H\right)^2}{2\left(1 - \underline{a}^H\right)}\right) - \left(\frac{\overline{b}^L - \overline{b}^H}{2}\right) \\
&= \frac{\left(1 + \underline{a}^L\right)}{2} - \frac{\left(1 + \underline{a}^H\right)}{2} - \left(\frac{\overline{b}^L - \overline{b}^H}{2}\right) \\
&= \frac{\underline{a}^L - \underline{a}^H - \overline{b}^L + \overline{b}^H}{2} = \frac{\underline{a}^L - \overline{b}^L - \left(\underline{a}^H - \overline{b}^H\right)}{2}
\end{aligned}$$

which is the average between the bid ask spreads charged by the most efficient dealer in the economies with low and high cost technology. Substituting out from the equilibrium values for $\underline{a}^L, \overline{b}^L, \underline{a}^H, \overline{b}^H$ explicitly, the difference in average bid ask spreads is

$$\begin{aligned}
s^L(0, \varepsilon) - s^H(k_m, 0) &= \frac{1}{2}\left\{\frac{1 - \lambda + 2\lambda_\varepsilon^2}{2\left(1 - \lambda + \lambda_\varepsilon^2\right)} - \left(\lambda_\varepsilon \frac{(1 - \lambda)}{2\left(1 - \lambda + \lambda_\varepsilon^2\right)}\right)\right\} \\
&\quad - \frac{1}{2}\left(\frac{1 + 2(1 - \lambda) + k_m}{2(2 - \lambda)} - \frac{(1 - \lambda)(1 - k_m)}{2(2 - \lambda)}\right) \\
&= \frac{1}{2}\left\{\frac{(1 - \lambda) + 2\lambda_\varepsilon^2 - \lambda_\varepsilon(1 - \lambda)}{2\left(1 - \lambda + \lambda_\varepsilon^2\right)} - \frac{(1 + k_m)}{2}\right\}
\end{aligned}$$

Then, the average bid ask spread is lower in the economy without insurance but with the high cost technology if and only if:

$$\begin{aligned}\frac{(1-\lambda)+2\lambda_\varepsilon^2-\lambda_\varepsilon(1-\lambda)-(1-\lambda+\lambda_\varepsilon^2)}{(1-\lambda+\lambda_\varepsilon^2)} &< k_m\\ \frac{\lambda_\varepsilon^2-\lambda_\varepsilon(1-\lambda)}{(1-\lambda+\lambda_\varepsilon^2)} &< k_m\end{aligned}$$

□

D Central clearing implementation

We consider the simplest implementation of central clearing in the model of Section 4, which we modify simply by introducing a continuum $[0,1]$ of dealers for each type k. Notice that this modification leaves all the derivations and results in the previous sections unchanged. If all dealers clear their transactions centrally via a Central Counterparty (CCP), then they must post collateral in the form of (margins, default fund contributions, and) default assessment. [18] Because the settlement shock ε is i.i.d. across dealers in each period, then it is i.i.d. also across the $[0,1]$ continuum of dealers of a given type k.

Suppose that all dealers are insured against the settlement shock ε, as we later verify. Therefore, they post bid and ask prices under the expectation that they face no such shock and that only a fraction λ of buyers will fail to settle their buy orders. This is equivalent to a version of the model with no settlement risk, described in Section 3, with the only difference being $\lambda \neq 0$. Let $a(k), b(k)$ denote the ask and bid prices posted by dealers of type k, and let $D(a(k)), S(b(k))$ denote the demand and supply for the asset which dealers of type k face from buyers and sellers respectively. Consistently with the analysis carried out in Section 3, each dealer chooses $a(k), b(k)$ to maximize expected profits $\Pi(k)$ subject to the feasibility constraint $(1-\lambda)D(a(k)) = S(b(k))$. As in Section 3, if a dealer posted ask and bid prices a, b then its demand and supply, at the stage where buy and sell orders are placed, satisfy

[18] For a description of the risk management practices of CCPs see BIS, and, for examples of default waterfall in CCPs currently operating in OTC markets, see ISDA [2013], EUR [2017], ICE [2017], DTC [2017] and LCH [2017]. For a rigorous modeling of the economic functions of a CCP, among which insurance against counterparty risk, see Acharya and Bisin [2014], Koeppl and Monnet [2013], Koeppl et al. [2012], Biais et al. [2016] and Biais et al. [2012].

$D(a) = \frac{(1-a)}{N}$ and $S(b) = \frac{b}{N}$. Then, a dealer with transaction cost k chooses a, b to solve:

$$\Pi(k) = \max_{a,b} \quad \{a(1-\lambda)D(a) - (b+k)S(b)\} \tag{41}$$

$$\text{s.t.} \quad (1-\lambda)D(a) \leq S(b) \tag{42}$$

As in Section 3, the feasibility constraint yields $b = (1-\lambda)(1-a)$, which substituted back into the objective function yields:

$$\Pi(k) = \max_a \{a(2-\lambda) - k - (1-\lambda)\} (1-\lambda)D(a) \tag{43}$$

Substituting out for $D(a)$ and taking first order conditions yields:

$$a(k) = \frac{3+k-2\lambda}{2(2-\lambda)} \tag{44}$$

$$b(k) = \frac{(1-\lambda)(1-k)}{2(2-\lambda)} \tag{45}$$

After the settlement shock is realized, a measure $\frac{1}{2}$ of dealers of type k receives shock $s = -1$ and its effective demand for the asset is $(1-\lambda-\varepsilon)D(a(k))$. Let $S_1(k)$ denote the set of such dealers. Analogously, a measure $\frac{1}{2}$ of dealers of type k receives shock $s = 1$ and its effective demand for the asset is: $(1-\lambda+\varepsilon)D(a(k))$. Let $S_2(k)$ denote the set of such dealers. Finally, let $d^k(s)$ denote the default assessment of dealer k towards the CCP when its idiosyncratic state is s, where $d^k : \{-1, +1\} \to \mathbb{R}$. [19] Under the rules of a CCP default waterfall, clearing members must contribute financial resources, so-called assessments, when necessary to avoid the CCP's default on any given position. Thus, a dealer $i \in S_1(k)$ faces effective demand $(1-\lambda-\varepsilon)D(a(k))$, but purchased $S(b(k)) = (1-\lambda)D(a(k))$ assets from sellers. As a consequence, such a dealer holds an excess of $\varepsilon D(a(k))$ assets purchased from sellers and unsold to buyers. On the contrary, a dealer $j \in S_2(k)$ faces effective demand $(1-\lambda+\varepsilon)D(a(k))$, but purchased only $S(b(k)) = (1-\lambda)D(a(k))$ assets from sellers. As a

[19] Notice that, because we have no collateral in the model, the default fund contribution by each CCP member takes place ex-post. In this respect the contribution is more similar to a default assessment, which usually occur after the margins and default fund contributions of defaulting and non-defaulting members have already been utilized.

consequence, such a dealer does not hold a sufficient inventory of assets to serve all of its buyers, and is short $\varepsilon D(a(k))$ assets. The CCP assessment mechanism can then insure both dealers ex-ante, by charging dealers $i \in S_1(k)$ an assessment $d^k(-1) = \varepsilon D(a(k))$ and dealers $j \in S_2(k)$ an assessment $d^k(+1) = -\varepsilon D(a(k))$. In other words, the former dealer makes a transfer of $\varepsilon D(a(k))$ assets to the latter. This process is described in Figure 4.

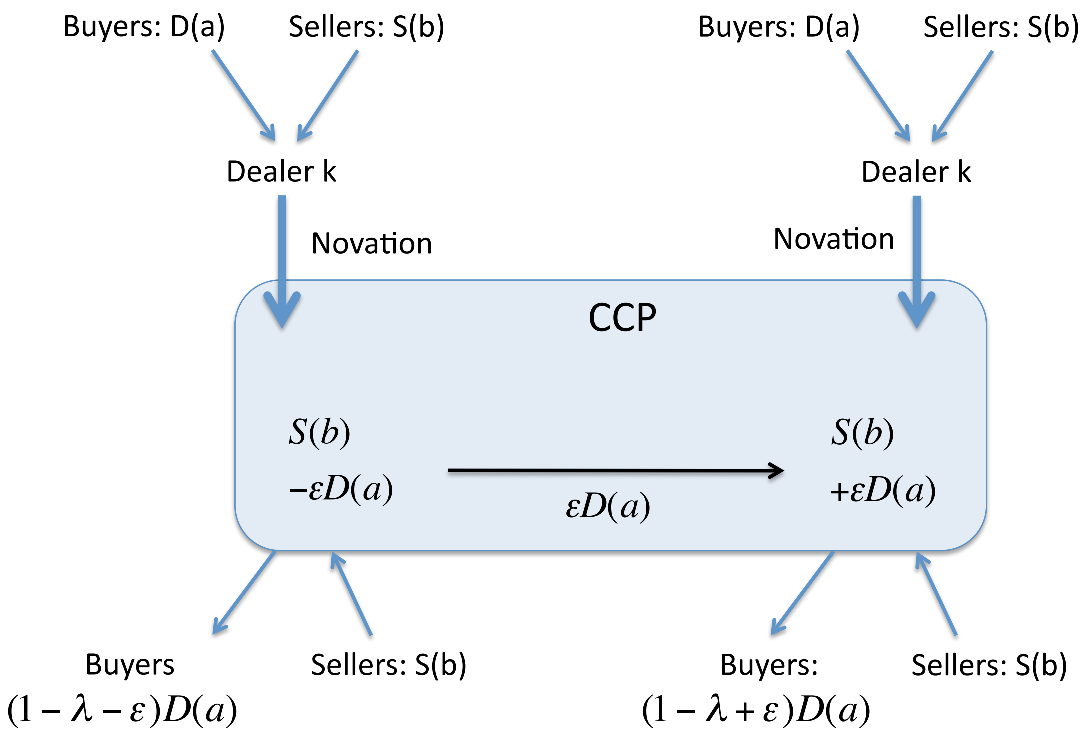

Figure 4: Implementation of central clearing in the model

In order to verify that (44) and (45) are indeed dealers' optimal response to the default assessment rule d^k, notice that dealers' k feasibility constraint in state $s = -1$ and $s = 1$ are, respecticely:

$$\begin{aligned}(1-\lambda-\varepsilon)D(a) &= S(b) - d^k(-1) = S(b) - \varepsilon D(a)\\(1-\lambda+\varepsilon)D(a) &= S(b) - d^k(+1) = S(b) + \varepsilon D(a)\end{aligned}$$

Notice that the feasibility constraints (6) boil down to (42) independent of the value of the settlement shock s. Moreover, the objective function (5) is simply:

$$\Pi(k;\lambda,\varepsilon) = \mathbb{E}_s\{a(1-\lambda+s\varepsilon)D(a)-(b+k)S(b)\} \tag{46}$$

Since $\mathbb{E}_s s = 0$ then the objective function of dealers k is simply (41). Therefore, the solution to dealers' k maximization problem yields (44) and (45).

E Risk aversion

We now consider the case where traders are risk averse in the following sense: The surplus from trade of a buyer is $x = v - a(k)$ whenever he accepts the bid price $a(k)$. Similarly the surplus from trade of a seller is $x = b(k) - v$. We assume that traders value the surplus from trade according to a CRRA utility function,

$$u(x) = \frac{(x+c)^{1-\sigma} - c^{1-\sigma}}{(1-\sigma)},$$

where $\sigma > 1$ and $c > 0$ is small. We need $c > 0$ so that traders prefer to trade than to exit the market without searching.[20] This specification implies that their decision to accept a bid or an ask price is the same as in the previous section. Therefore, the optimal bid and ask prices set by dealers (8)-(9) are unchanged. As a consequence, the least efficient dealer in operation is still $\bar{k}$ defined by (14). Also, the effect of settlement risk on the bid-ask prices is unchanged: Increased settlement risk makes entry less profitable so that the least efficient dealers exit the market. As a consequence, the distribution of ask-prices becomes more concentrated. While they face higher ask price, buyers face a lower dispersion of ask price. Since they are risk averse, they may prefer that dealer face a little more risk. Obviously, buyers face a trade-off as on one hand they face a higher average ask-price, but on the other hand, the distribution of ask price is more compressed.

It is tedious to compute the overall buyers' welfare with $c > 0$ and we do so in the

[20]This is the case if $\sigma > 1$ as $x^{1-\sigma}/(1-\sigma) < 0$ for all $x \geq 0$, and this affects the decision of traders to accept or reject an offer.

Appendix where we show that with $c > 0$,

$$U_c = \frac{(1-\lambda)}{(1-\sigma)} \left\{ \frac{(1-a(0)+c)^{3-\sigma} - c^{3-\sigma}}{(1-a(0))(2-\sigma)(3-\sigma)} - \frac{c^{1-\sigma}}{2}(1-a(0)) - \frac{c^{2-\sigma}}{2-\sigma} \right\}$$

Hence, we obtain

$$\frac{\partial U_c}{\partial \varepsilon} = \frac{(1-\lambda)}{(1-\sigma)} \left\{ -\frac{(1-a(0)+c)^{2-\sigma}}{(1-a(0))(2-\sigma)} + \frac{(1-a(0)+c)^{3-\sigma} - c^{3-\sigma}}{(1-a(0))^2(2-\sigma)(3-\sigma)} + \frac{c^{1-\sigma}}{2} \right\} \frac{\partial a(0)}{\partial \varepsilon}$$

Computation with different values for σ reveals that the payoff of consumers is always decreasing with an increasing in settlement risk. Therefore, concavity of the buyer's payoff function is not enough to generate the desirability of settlement risk. We turn next to different distribution of the dealers' cost.

E.1 Distribution function for dealers transaction cost

In this section of the paper we assume that dealers are distributed according to a beta probability distribution $f(k;\alpha,\beta) = \frac{\alpha k^{\alpha-1}(1-k)^{\beta-1}}{\mathcal{B}(\alpha,\beta)}$ with support $[0,1]$. Let $\beta = 1$ so that $\mathcal{B}(\alpha,\beta) = 1$. Then the cdf associated with it is

$$F(k) = \int_0^k \alpha s^{\alpha-1} ds = k^\alpha$$

Now, because only $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon} < 1$ are active, then[21]

$$F_{\overline{k}}(k) = \frac{k^\alpha}{\overline{k}^\alpha}$$

and the probability distribution function is then simply $f_{\overline{k}}(k) = \alpha \frac{k^{\alpha-1}}{\overline{k}^\alpha}$.

Notice that ask prices are an affine transformation of the dealer's cost of the form $a(k) =$

[21]Or, similarly, from $F(k) = k^\alpha$ we have that the truncated distribution $F_{\overline{k}}(k) = \Pr\left(s \le k \mid s \le \overline{k}\right) = \frac{\Pr\left(s \le k \cap s \le \overline{k}\right)}{\Pr\left(s \le \overline{k}\right)} = \frac{F(k)}{F\left(\overline{k}\right)}$.

$a(0)+\xi k$ where $a(\bar{k})=1$ and $\xi=\frac{\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, then the cdf of $a\left(k\right)$ is derived from $F_{\overline{k}}\left(k\right)$:

$$\begin{aligned} F_a\left(\hat{a}\right) &= \frac{\left(\frac{\hat{a}-a(0)}{\xi}\right)^\alpha}{\overline{k}^\alpha} \\ f_a\left(a\right) &= \frac{1}{\xi}f_{\bar{k}}\left(\frac{a-a(0)}{\xi}\right) \end{aligned}$$

Similarly for the bid price

$$b\left(k\right) = b\left(0\right)-\lambda_\varepsilon\xi k$$

And

$$\begin{aligned} F_b\left(\hat{b}\right) &= 1-\frac{\left(\frac{b(0)-\hat{b}}{\lambda_\varepsilon\xi}\right)^\alpha}{\overline{k}^\alpha} \\ f_b\left(b\right) &= \frac{1}{\lambda_\varepsilon\xi}f_{\bar{k}}\left(\frac{b\left(0\right)-b}{\lambda_\varepsilon\xi}\right) \end{aligned}$$

E.2 Consumers' surplus

Then consumers' surplus (with linear preferences), using integration by parts, is:

$$\begin{aligned} S_c &= \int_{a(0)}^{1}\left[\int_{a(0)}^{v}(v-a)f_a(a)da\right]dv \\ &= \frac{\left(1-a\left(0\right)\right)^{\alpha+2}}{\xi^\alpha\overline{k}^\alpha\left(\alpha+1\right)\left(\alpha+2\right)} \end{aligned}$$

Using $a(k)=1-\frac{1-\lambda-k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, $\xi=\frac{\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\overline{k}=\frac{1-\lambda}{\lambda_\varepsilon}$ we then have:

$$S_c = \frac{\left(\frac{1-\lambda}{2(1-\lambda+\lambda_\varepsilon^2)}\right)^2}{\left(\alpha+1\right)\left(\alpha+2\right)}$$

which is decreasing in ε. Also notice that the smaller α is the faster S_c decreases in ε.

E.3 Producers' surplus

Similarly for producers' surplus, using integration by parts:

$$
\begin{aligned}
S_p &= \int_0^{b(0)} \left[\int_v^{b(0)} (b-v) f_b(b) db \right] dv \\
&= \frac{b\left(0\right)^{\alpha+2}}{\left(\alpha+1\right)\left(\alpha+2\right)\left(\lambda_\varepsilon \xi \overline{k}\right)^{\alpha}}
\end{aligned}
$$

Using $b(k) = \lambda_\varepsilon \frac{1-\lambda-k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ we then have:

$$
S_p = \lambda_\varepsilon^2 S_c
$$

which is increasing[22] in ε if and only if $\varepsilon \in [0, \overline{\varepsilon}]$ (where $\overline{\varepsilon} = -\left(1-\lambda\right) + \sqrt{1-\lambda}$ as defined above). Also notice that the smaller α is the faster S_p increases in ε.

E.4 Dealers' surplus

For dealers let us rewrite the expected demand and supply faced in their decision problem:

$$
D\left(a\right) = \int_a^{\overline{r}_c} \tilde{h}\left(r\right) dr
$$

[22]Where

$$
\begin{aligned}
\frac{\partial S_p}{\partial \varepsilon} &= \frac{\left(1-\lambda\right)^2}{4\left(\alpha+1\right)\left(\alpha+2\right)} \frac{\partial \left(\frac{\lambda_\varepsilon}{(1-\lambda+\lambda_\varepsilon^2)}\right)^2}{\partial \varepsilon} \\
&= \frac{\left(1-\lambda\right)^2}{4\left(\alpha+1\right)\left(\alpha+2\right)} \frac{2\lambda_\varepsilon}{\left(1-\lambda+\lambda_\varepsilon^2\right)} \left(\frac{1-\lambda+\lambda_\varepsilon^2-2\lambda_\varepsilon^2}{\left(1-\lambda+\lambda_\varepsilon^2\right)^2}\right) \\
&= \frac{\lambda_\varepsilon\left(1-\lambda\right)^2}{2\left(\alpha+1\right)\left(\alpha+2\right)} \frac{\left(1-\lambda-\lambda_\varepsilon^2\right)}{\left(1-\lambda+\lambda_\varepsilon^2\right)^3}
\end{aligned}
$$

which is always strictly positive if and only if ε is such that $1-\lambda-\lambda_\varepsilon^2 > 0$.

where $\tilde{h}(r)$ is the conditional probability density of consumers' reservation prices among the fraction $1-\underline{v}_c$ who chose to participate in the dealers' market. Therefore, $\tilde{h}(r)$ is derived as follows: the reservation price of a consumer with valuation v, denoted $r_c(v)$, is simply that specific consumer's valuation:

$$r_c(v) = v$$

Now, $v \sim U[\underline{v}_c, 1]$ therefore

$$\begin{aligned}\Pr(r_c(v) \leq r) &= \Pr(v \leq r) \\ &= \frac{r-\underline{v}_c}{1-\underline{v}_c}\end{aligned}$$

and the probability density function associated with it is simply $h(r) = \frac{1}{1-\underline{v}_c}$. Then the per dealer k density of consumers is $(1-\underline{v}_c) f_{\overline{k}}(k) h(r)$. So that the mass of consumers who place an order when the ask price they face is a (i.e. demand faced by a dealer who posts ask price a if his type is k -because here the mass of consumers that contact him is a function of k) is simply

$$\begin{aligned}D(a(k)) &= \int_{a(k)}^{\overline{r}_c} (1-\underline{v}_c) f_{\overline{k}}(k) h(r) dr \\ &= (1-a(k)) f_{\overline{k}}(k)\end{aligned}$$

And similarly for the supply:

$$S(b(k)) = b(k) f_{\overline{k}}(k)$$

For dealers, we also need to take into account the constraint of meeting demand period by period, so that substituting the expected demand and supply per dealer k into the objective function of a dealer we have, as before, that expected profits of dealer k with the optimal choice of a, are

$$\begin{aligned}\pi(k;\lambda,\varepsilon) &= f_{\overline{k}}(k)\left\{a(k)(1-\lambda)-\left[\lambda_\varepsilon(1-a(k))+k\right]\lambda_\varepsilon\right\}(1-a(k)) \\ &= \alpha\frac{k^{\alpha-1}}{\overline{k}^{\alpha}}\frac{(1-\lambda-k\lambda_\varepsilon)^2}{4(1-\lambda+\lambda_\varepsilon^2)}\end{aligned}$$

Then aggregate dealers' surplus is given by the total discounted profits of all dealers participating in the dealer market are:

$$\begin{aligned}S_d(\varepsilon) &= \int_0^{\bar{k}}\Pi(k;\lambda,\varepsilon)dk \\ &= \frac{1}{4(\alpha+1)(1-\lambda+\lambda_\varepsilon^2)}\left\{\left(1-\lambda-\lambda_\varepsilon\bar{k}\right)\left[(\alpha+1)(1-\lambda)+\left[2\lambda_\varepsilon-(\alpha+1)\lambda_\varepsilon\right]\bar{k}\right]+\frac{2\lambda_\varepsilon^2\bar{k}^2}{(\alpha+2)}\right\}\end{aligned}$$

And using $\overline{k}=\frac{1-\lambda}{\lambda_\varepsilon}$ we then have:

$$S_d = \frac{(1-\lambda)^2}{2(\alpha+1)(\alpha+2)(1-\lambda+\lambda_\varepsilon^2)}$$

Also notice that the smaller α is the faster S_d decreases in ε.

Overall we have the following result:

Claim 1. S_d decreases in ε. The smaller α is the larger is the decrease in S_d. S_p increases in ε, for $\varepsilon\in[0,\overline{\varepsilon}]$, and decreases in ε, for $\varepsilon\in[\overline{\varepsilon},\lambda]$. The smaller α is the larger is the increase (decrease) in S_d. S_c is decreasing in ε. The smaller α is the faster S_c decreases in ε.

E.5 Total welfare

Summing up consumers', producers' and dealers' welfare we have:

$$\begin{aligned}W &= S_c+S_p+S_d \\ &= \frac{(1-\lambda)^2(3+3\lambda_\varepsilon^2-2\lambda)}{4(\alpha+1)(\alpha+2)(1-\lambda+\lambda_\varepsilon^2)^2}\end{aligned}$$

And:

$$\frac{\partial W}{\partial \varepsilon} = \frac{(1-\lambda)^2 \lambda_\varepsilon}{2(\alpha+1)(\alpha+2)} \frac{(\lambda - 3(1+\lambda_\varepsilon^2))}{(1-\lambda+\lambda_\varepsilon^2)^3}$$

which is always negative since

$$\lambda - 3\left(1+\lambda_\varepsilon^2\right) < 0$$

Claim 2. Total welfare is always decreasing in ε regardless of the value of α.

E.6 Different parameters for beta distribution

In this section of the paper we assume that dealers are distributed according to a beta probability distribution $f(k;\alpha,\beta) = \frac{\beta k^{\alpha-1}(1-k)^{\beta-1}}{\mathcal{B}(\alpha,\beta)}$ with support $[0,1]$. Let $\alpha = 1$ so that $f(k;\alpha,\beta) = \beta(1-k)^{\beta-1}$ and the cdf associated with it is

$$F(k) = 1-(1-k)^\beta$$

Now, because only $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon} < 1$ are active, then[23]

$$F_{\overline{k}}(k) = \frac{1-(1-k)^\beta}{1-\left(1-\overline{k}\right)^\beta}$$

and the probability distribution function is then simply $f_{\overline{k}}(k) = \beta\frac{(1-k)^{\beta-1}}{1-\left(1-\overline{k}\right)^\beta}$.

Notice that ask prices are an affine transformation of the dealer's cost of the form $a(k) = a(0) + \xi k$ where $a(\bar{k}) = 1$ and $\xi = \frac{\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, then the cdf of $a(k)$ is derived from $F_{\overline{k}}(k)$:

$$F_a(\hat{a}) = \frac{1-\left(1-\frac{\hat{a}-a(0)}{\xi}\right)^\beta}{1-\left(1-\overline{k}\right)^\beta}$$

[23] Or, similarly, from $F(k) = k^\alpha$ we have that the truncated distribution $F_{\overline{k}}(k) = \Pr\left(s \le k \mid s \le \overline{k}\right) = \frac{\Pr\left(s\le k \cap s \le \overline{k}\right)}{\Pr\left(s \le \overline{k}\right)} = \frac{F(k)}{F\left(\overline{k}\right)}$.

$$f_a(a) = \frac{1}{\xi} f_{\bar{k}} \left(\frac{a - a(0)}{\xi} \right)$$

Similarly for the bid price

$$b(k) = b(0) - \lambda_\varepsilon \xi k$$

And

$$\begin{aligned} F_b\left(\hat{b}\right) &= \frac{\left(1 - \frac{b(0)-\hat{b}}{\lambda_\varepsilon \xi}\right)^\beta - \left(1 - \overline{k}\right)^\beta}{1 - \left(1 - \overline{k}\right)^\beta} \\ f_b(b) &= \frac{1}{\lambda_\varepsilon \xi} f_{\bar{k}} \left(\frac{b(0) - b}{\lambda_\varepsilon \xi} \right) \end{aligned}$$

E.6.1 Consumers' surplus

Then consumers' surplus (with linear preferences), using integration by parts, is:

$$\begin{aligned} S_c &= \int_{a(0)}^{1} \left[\int_{a(0)}^{v} (v-a) f_a(a) da \right] dv \\ &= \frac{1 - a(0)}{\left[1 - \left(1 - \overline{k}\right)^\beta\right]} \left(\frac{1 - a(0)}{2} - \frac{\xi}{\beta + 1} \right) - \frac{\frac{\xi}{\beta+1} \frac{\xi}{\beta+2}}{\left[1 - \left(1 - \overline{k}\right)^\beta\right]} \left[\left(1 - \frac{1 - a(0)}{\xi}\right)^{\beta+2} - 1 \right] \end{aligned}$$

And using $a(k) = 1 - \frac{1-\lambda-k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, $\xi = \frac{\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ we then have that for $\lambda = 0.3, \beta = 0.2$ consumers' surplus as a function of ε is increasing for small values of ε, as Figure 5 shows.

Let ε^* denote the threshold such that $\forall \varepsilon \leq \varepsilon^*$ we have that $\frac{\partial S_c}{\partial \varepsilon} > 0$ and $\forall \varepsilon > \varepsilon^*$ we have that $\frac{\partial S_c}{\partial \varepsilon} < 0$. Then as $\beta > 0$ decreases we have that ε^* increases. Also, for the same value of β, ε^* is decreasing in λ. Figure 6 shows S_c as a function of ε for $\lambda = 0.1, \beta = 0.2$.

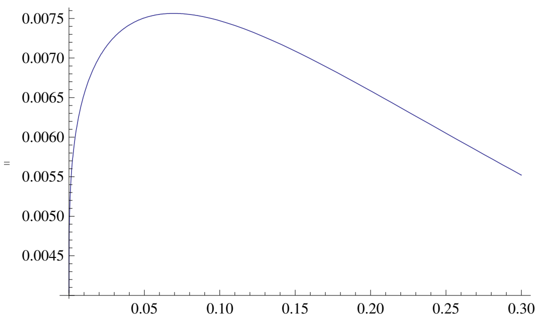

Figure 5: Consumers' surplus as a function of ε: $\lambda = 0.3, \beta = 0.2$

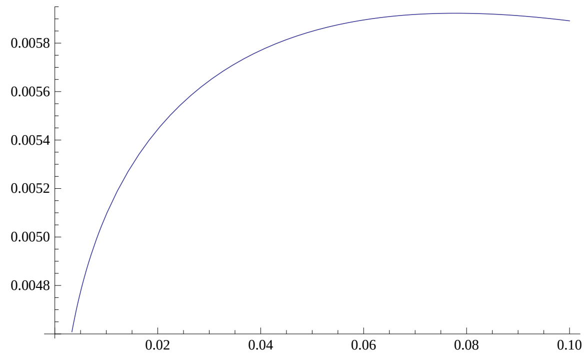

Figure 6: Consumers' surplus as a function of ε: $\lambda = 0.1, \beta = 0.2$

And substituting out $a(k) = 1 - \frac{1-\lambda-k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, $\xi = \frac{\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$ we then have that:

$$S_c = \frac{(\beta+2)(1-\lambda)(\beta+1)(1-\lambda) - 2\lambda_\varepsilon^2\left(1-\frac{1-\lambda}{\lambda_\varepsilon}\right)^{\beta+2} + 2\lambda_\varepsilon\left(\lambda_\varepsilon - (\beta+2)(1-\lambda)\right)}{8(\beta+1)(1-\lambda+\lambda_\varepsilon^2)^2\left[1-\left(1-\frac{1-\lambda}{\lambda_\varepsilon}\right)^{\beta}\right](\beta+2)}$$

The whole positive effect of ε comes from $\left[1-\left(1-\frac{1-\lambda}{\lambda_\varepsilon}\right)^{\beta}\right]$ at the denominator which is coming form $\overline{k}$ through the distribution of ask prices. Figure 7 shows the pdf of the ask price, $f_a(a)$, for $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.02$. Notice that when ε increases, the mass on every surviving dealer increases. Figure 8 shows the pdf of the ask price, $f_a(a)$, for $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.05$. Figure 9 shows the pdf of the ask price, $f_a(a)$, for $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.2$.

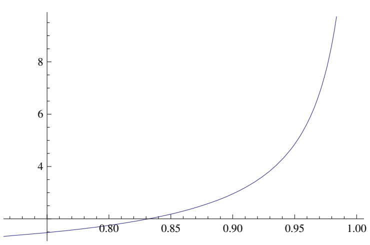

Figure 7: $f_a(a)$: $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.02$

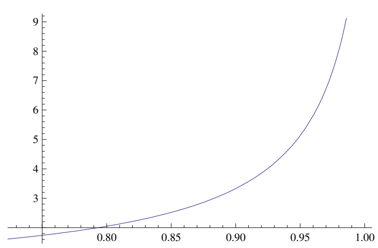

Figure 8: $f_a(a)$: $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.05$

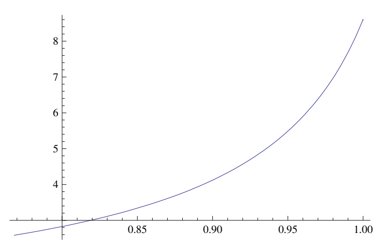

Figure 9: $f_a(a)$: $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.2$

E.6.2 Producers' surplus

Similarly for producers' surplus, using integration by parts:

$$
\begin{aligned}
S_p &= \int_0^{b(0)} \left[\int_v^{b(0)} (b-v) f_b(b) db \right] dv \\
&= \frac{1}{\left[1-\left(1-\overline{k}\right)^{\beta}\right]} \left\{ \frac{b(0)^2}{2} - \frac{\lambda_\varepsilon \xi}{\beta+1} b(0) + \frac{\lambda_\varepsilon \xi}{\beta+1} \frac{\lambda_\varepsilon \xi}{\beta+2} \left[1 - \left(1 - \frac{b\left(0\right)}{\lambda_\varepsilon \xi} \right)^{\beta+2} \right] \right\}
\end{aligned}
$$

Using $b(k) = \lambda_\varepsilon \frac{1-\lambda-k\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$, $\xi = \frac{\lambda_\varepsilon}{2(1-\lambda+\lambda_\varepsilon^2)}$ and $\overline{k} = \frac{1-\lambda}{\lambda_\varepsilon}$,we then have:

$$
S_p = \frac{\lambda_\varepsilon^2}{\left[1-\left(1-\frac{1-\lambda}{\lambda_\varepsilon}\right)^{\beta}\right]} \frac{\left(\beta+1\right)\left(1-\lambda\right)^2 + \frac{2\lambda_\varepsilon^2}{(\beta+2)} \left(1-\left(1-\frac{1-\lambda}{\lambda_\varepsilon}\right)^{\beta+2}\right) - 2\lambda_\varepsilon \left(1-\lambda\right)}{4\left(\beta+1\right)\left(1-\lambda+\lambda_\varepsilon^2\right)^2}
$$

Interestingly, also the producers' surplus is decreasing in ε for large values of β: for example for $\beta = 2, \lambda = 0.3$ it is decreasing, but for $\beta = 1, \lambda = 0.3$ it is hump shaped with a threshold ε^*such that $\forall \varepsilon \leq \varepsilon^*$we have that $\frac{\partial S_p}{\partial \varepsilon} > 0$ and $\forall \varepsilon > \varepsilon^*$we have that $\frac{\partial S_p}{\partial \varepsilon} < 0$. As in the consumers' surplus case, as $\beta > 0$ decreases we have that ε^* increases. Figure 10 shows producers' surplus, S_p, as a function of ε when $\beta = 0.7, \lambda = 0.1$. Notice that for sufficiently small values of λ producers' surplus is strictly increasing in ε, while for sufficiently large values of λ, as long as β is small enough, then producers' surplus is hump shaped as a function of ε. Figure 11 shows producers' surplus S_p as a function of ε when $\beta = 0.7, \lambda = 0.9$. In order to gain insight on what is going on with the distribution of bid prices, Figure 12 shows the pdf of the bid price, $f_b\left(b\right)$, for $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.02$.

Therefore there is a lot of mass on inefficient dealers so that when they exit all that mass gets thrown onto more efficient dealers: recall that more efficient dealers are the ones who charge the highest (lowest) bid (ask) price because they are the only ones who can afford to do so. Therefore the above picture means that few dealers (the efficient ones) charge the highest bid prices, whereas many dealers (the inefficient ones) charge the lowest bid prices.

Notice that when ε increases, the mass on bid prices offered by very efficient dealers

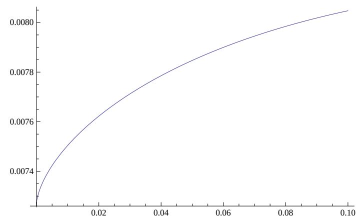

Figure 10: $S_p(\varepsilon)$: $\beta = 0.7, \lambda = 0.1$

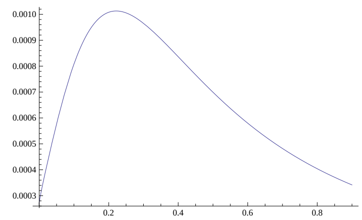

Figure 11: $S_p(\varepsilon)$: $\beta = 0.7, \lambda = 0.9$

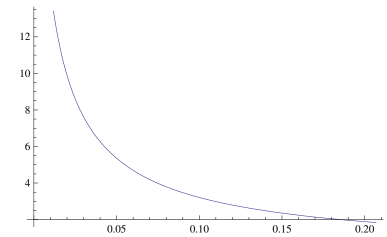

Figure 12: $f_b(b)$: $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.02$

increases. Figure 13 (red) shows the pdf of $f_b(b)$ for $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.02$ and Figure 14 (green) for $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.05$.

Notice that $\underline{b} = 0$ is unchanged because it is the bid price quoted by the marginal operating dealer (which is making zero profits); however $\overline{b}$ increases with ε because it is the bid price quoted by the most efficient dealer whose demand and supply change as ε increases because there are less dealers who are active (since $\overline{k}$ decreases). Therefore the most efficient dealer is more likely to get a random call by a buyer and a seller ($f_{\overline{k}}(k = 0)$ increases) and he is efficient enough that it is profitable for him to increase the bid price and serve a larger share of the market.

E.6.3 Dealers' surplus

If we take into account that expected demand and supply are $D(a) = (1 - a(k)) f_{\overline{k}}(k)$ and $S(b(k)) = b(k) f_{\overline{k}}(k)$then expected profits are $\pi(k; \lambda, \varepsilon) = \beta \frac{(1-k)^{\beta-1}}{1-\left(1-\overline{k}\right)^{\beta}} \frac{(1-\lambda-k\lambda_{\varepsilon})^2}{4(1-\lambda+\lambda_{\varepsilon}^2)}$. Either way we know that the calculation of aggregate dealers' surplus is the same regardless of which

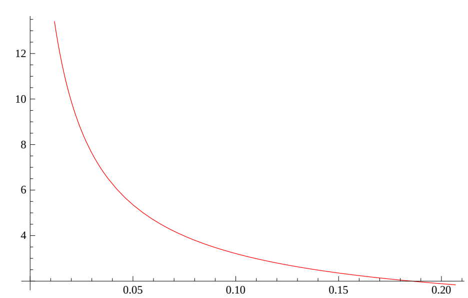

Figure 13: $f_b(b)$: $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.02$

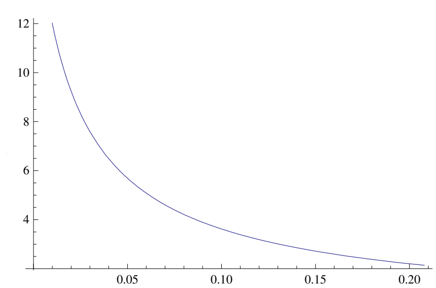

Figure 14: $f_b(b)$: $\beta = 0.2, \lambda = 0.3, \varepsilon = 0.05$

interpretation we give (matching or probability). Therefore aggregate dealers' surplus is:

$$\begin{aligned} S_d(\varepsilon) &= \int_0^{\bar{k}} \Pi(k; \lambda, \varepsilon) dk \\ &= \frac{(1-\lambda)}{\left(1-\left(1-\bar{k}\right)^{\beta}\right) 4\left(1-\lambda+\lambda_{\varepsilon}^2\right)} \left\{(1-\lambda) - \frac{2\lambda_{\varepsilon}}{\beta+1} + \left(2\lambda_{\varepsilon}\bar{k} - (1-\lambda) + \frac{2\lambda_{\varepsilon}}{\beta+1}\left(1-\bar{k}\right)\right)\left(1-\bar{k}\right)^{\beta}\right\} + \\ &\quad + \frac{\lambda_{\varepsilon}^2}{\left(1-\left(1-\bar{k}\right)^{\beta}\right) 4\left(1-\lambda+\lambda_{\varepsilon}^2\right)} \left\{\frac{2}{(\beta+1)(\beta+2)} - \left(1-\bar{k}\right)^{\beta}\left[\bar{k}^2 + \frac{2\left(\left(1-\bar{k}\right)^2 + (\beta+2)\left(1-\bar{k}\right)\bar{k}\right)}{(\beta+1)(\beta+2)}\right]\right\} \end{aligned}$$

Dealers' surplus for a given λ is inverse U-shaped in ε: in general the smaller λ the larger the value of β^*, where $\beta^* = \left\{\beta > 0 : \frac{\partial S_d}{\partial \varepsilon} > 0, \forall \beta < \beta^*\right\}$. For a given λ, as we increase β the peak of the inverse U shaped function is reached at a value $\hat{\varepsilon} < 0$; analogously for β small the peak of the inverse U shaped function is reached at a value $\hat{\varepsilon} > \lambda$, therefore in these two polar cases we have that dealers' surplus is either decreasing, increasing or hump-shaped in any feasible value of $\varepsilon \in [0, \lambda]$, as we can see from the figures below.

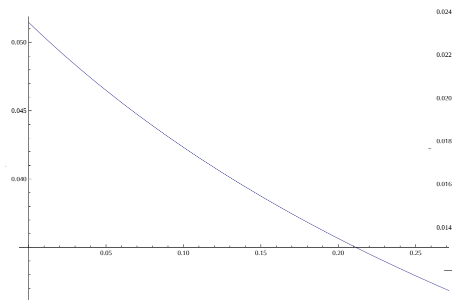

Figure 15: $S_d(\varepsilon)$: $\beta = 2, \lambda = 0.3$

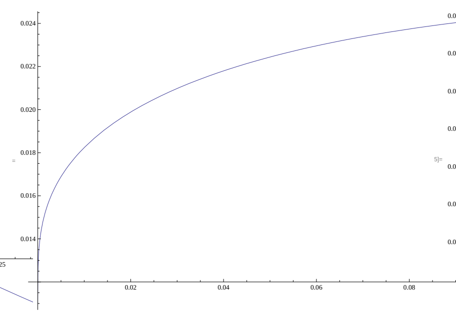

Figure 16: $S_d(\varepsilon)$: $\beta = 0.2, \lambda = 0.1$

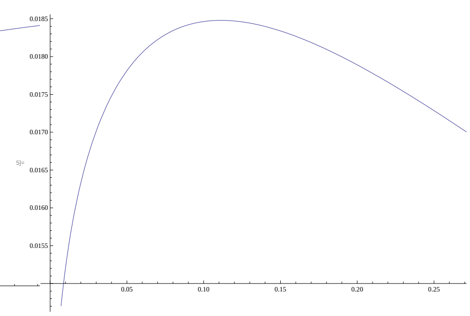

Figure 17: $S_d(\varepsilon)$: $\beta = 0.2, \lambda = 0.3$