

Inefficiency in a Simple Model of Production and Bilateral Trade

Zachary Bethune, Bruno Sultanum, and Nicholas Trachter

Many assets, both real and financial, are traded in secondary over-the-counter (OTC) markets after their initial issuance (e.g., real estate, municipal bonds, Treasuries, asset-backed securities, etc.). The apparent importance of OTC markets has prompted an explosion of research on the topic. There is a large literature studying OTC markets with no ex-ante investment (see, Duffie et al. 2005, Duffie et al. 2007, Lagos and Rocheteau 2009, and Hugonnier et al. 2014, among others). A new and growing literature studies how the market structure affects ex-ante decisions. In Farboodi et al. (2017), the market structure interacts with the investors' incentives to invest in negotiation skills (which affects bargaining power). In Farboodi et al. (2016), the market structure interacts with the incentives to invest on the speed at which an investor will meet others in the secondary market. In Bethune et al. (2017), the market structure affects the incentives to produce new assets and to introduce them into the market. In these last three papers, the equilibrium is constrained-inefficient. Our goal in this paper is to study a source of inefficiency that stems from these models in a much simpler setup. To that end, we will work with a simplified version of Bethune et al. (2017).

We study a simple model of production and trade in OTC markets. The model is meant to capture some important features of these markets. First, as in the seminal paper Duffie et al. (2005), trade is resolved in decentralized markets, where the buyer and seller in a trade

■ The views expressed herein are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. We thank the editor and the three referees for useful comments. Email: nicholas.trachter@rich.frb.org.

meeting bargain over the terms of trade. Second, agents face uncertainty over whether their potential trade counterparty wants to or can trade with her. Third, as in Bethune et al. (2017), we endogenize the asset level by allowing agents to produce assets at some cost. Unlike these papers, and a large part of the literature on the topic, we do not include search frictions in the model. Instead, as in Babus and Kondor (2013) or Gofman (2014), we model the market with no search frictions.

The model is static. The economy is populated by two agents who produce and trade assets. The two agents differ in the level at which they value assets: one agent's asset valuation is ν_l , while the other agent's asset valuation is ν_h , with $\nu_h > \nu_l$. At the beginning of the only period in the model, agents are faced with an opportunity to produce an asset, but the production cost is random. Then, agents meet and trade. When trade happens, the trade surplus is divided among the trade participants using Nash bargaining, where θ denotes the bargaining power of the buyer. In the economy, the low-valuation agent acts as an intermediary: she produces the asset with the intention of offloading it during the meeting with the high-valuation agent. Thus, from now on, we refer to this agent as the intermediary. The high-valuation agent intends to obtain an asset. She can do so by either producing it herself or by enjoying the intermediation services provided by the low-valuation agent. We refer to this agent as the investor. We solve for the unique equilibrium in this economy, and we study its efficiency properties. The model is described and analyzed in Section 1.

We find that the equilibrium is never constrained-efficient in the sense that there is no way to split the surplus between the intermediary and the investor, i.e., no choice of $\theta \in [0, 1]$, that makes the equilibrium efficient. This result is consistent with Bethune et al. (2017), but it is different from what is found in many search models (for example, in the seminal work of Lagos and Wright [2005]) where there is always a θ such that the equilibrium is constrained-efficient. The fundamental difference lies in the fact that in these other environments there is a one-sided hold-up problem, while in ours—as in most of the recent literature on OTC markets—there is a double-sided hold-up problem.¹ The double-sided hold-up problem arises because both parties in a trade meeting have to make a costly investment. In order to have trade between the intermediary and investor, the intermediary has to produce the asset—pay a cost and gain the option of selling the asset, and thus she will be held up by the investor in the trade meeting. Likewise, for trade to happen, the investor has to refrain herself from

¹ Recent examples are Farboodi et al. (2016) and Farboodi et al. (2017).

producing the asset—to arrive at the trade meeting empty handed, she must give up the option of producing an asset, and thus she will also be held up by the intermediary in the trade meeting. The first hold-up problem could be solved by assigning all the bargaining power to the intermediary, i.e., $\theta = 0$ (given that the intermediary acts as the seller in the trade meeting). The second hold-up problem could be solved by assigning all the bargaining power to the investor, i.e., $\theta = 1$ (given that the investor acts as the buyer in the trade meeting). Both cannot be handled together. Thus, the equilibrium is inefficient. We study efficiency in Section 2.

Interestingly, the inefficiency implies that there can be over- or underinvestment in asset production, which translates into an over- or undersupply of assets in the economy. When θ is low, the inefficiency implies that there is overinvestment and thus an excess supply of assets. When θ is high, the inefficiency implies that there is underinvestment and thus an inefficiently low asset supply. We study the effect of the inefficiency on the asset supply in detail in Section 3.

Our result in this paper is the same result found in Bethune et al. (2017). In that paper, we study a dynamic model of OTC trade with production, where meetings among potential traders are subject to search frictions, as in Duffie et al. (2005) or Hugonnier et al. (2014). In this paper, we show that the inefficiency result does not stem from search frictions and its potential inefficiency related to the Hosios condition (Hosios 1990). In fact, there is no search in our model here. We also show that the inefficiency result vanishes when both intermediary and investor can perfectly forecast the actions of their trade counterparties through the equilibrium. We show this by studying the case where the upper bound of the production cost equals ν_h , which can be interpreted as the limit as the upper bound of the cost distribution approaches ν_h . It happens that, in this limit, the high-valuation agent invests for any θ and for any draw of production cost. This is perfectly forecastable by the intermediary, who therefore does not produce. We study this in Section 4. Finally, Section 5 concludes.

1. MODEL

We study a simple two-agent, one-period model of asset trading. Agents differ in their valuation of the asset. One agent values the asset at level ν_l , while the other agent values the asset at level ν_h , where $\nu_h > \nu_l$. Agents can hold either zero or one unit of the asset. At the beginning of the period, both agents are confronted with an opportunity to invest in producing an asset. The investment cost c is random, drawn from the cumulative distribution $G(c)$. This distribution is assumed to be

uniform in the support $[\underline{c}, \bar{c}]$ with $g(c) \equiv \partial G(c)/\partial c$. To keep things as easy as possible, we restrict attention to the case where $0 = \underline{c} = \nu_l < \nu_h < \bar{c} = 1$. The restriction implies that (i) agents with a low valuation of the asset only produce with the intention to sell and thus play the role of intermediaries, and (ii) because $\nu_h < \bar{c}$, we can abstract from potential kinks and discontinuities when analyzing optimal behavior.² Given (i), we refer to agents with low valuation as intermediaries and to agents with high valuation as investors. Finally, once production decisions are made, agents meet and trade. We restrict them to using a Nash bargaining protocol to resolve trade. We let $\theta \in [0, 1]$ denote the bargaining power of the buyer. We later argue that the investor plays the role of the buyer in the trade meeting, while the intermediary plays the role of the seller in the trade meeting. Thus, we can think of θ as the bargaining power of the investor.

We can solve this model using a backward-induction argument. First, conditional on an asset distribution among agents, we solve the problem where the two agents meet and trade. Then, with this solution at hand, we solve the investment stage of the model.

Trade. When agents meet and trade, the outcomes depend on who enters this stage holding an asset. When both investors are either holding or not holding an asset, there is trivially no trade. If the investor enters the trade stage holding an asset, there is no trade. This follows because the maximum willingness to pay for the asset by a intermediary is ν_l , which is lower than the cost of giving up the asset for the investor, which is ν_h . As a result, there is no positive surplus to be split by the two parties involved in trade. The only case where there will be trade is when the intermediary is holding an asset and the investor is not. Because $\nu_h > \nu_l$, the total surplus of trading is positive, and thus trade will happen. However, the way the surplus is split among the two traders depends on the way bargaining power is assigned to them. Let p denote the transfer, or price, that the seller—the intermediary—requires to transfer the asset to the buyer—the investor. The optimal price p solves

$$\max_p (p - \nu_l)^{1-\theta} (\nu_h - p)^\theta . \quad (1)$$

Notice that when agents trade: (i) the intermediary receives p in exchange for giving up the asset, which she values at ν_l , and (ii) the investor receives the asset, which she values at ν_h , and pays p . Operat-

² Because $\bar{c} > \nu_h$, there is always a draw for the production cost for which neither agent wants to produce. On the flip side, because $\underline{c} = 0 < \nu_l$, there is always a draw for the production cost for which both agents want to produce. These two observations, coupled with full support for the production cost in $[\underline{c}, \bar{c}]$, guarantees that the problem is smooth.

ing with the first-order condition of the problem presented in equation (1) provides an expression for the price,

$$p = \theta\nu_l + (1 - \theta)\nu_h , \quad (2)$$

which is simply a weighted average of the asset valuation of the seller and buyer. When the full bargaining power is assigned to the intermediary, i.e. $\theta = 0$, we obtain that $p = \nu_h$. That is, the intermediary obtains all the surplus of the trade. Likewise, when the full bargaining power is assigned to the investor, i.e. $\theta = 1$, we obtain that $p = \nu_l$. That is, the investor obtains all the surplus of the trade. In fact, given that $\partial p / \partial \theta = -(\nu_h - \nu_l) < 0$, the price decreases as the bargaining power of the investor increases, implying that a higher fraction of the total surplus of trade is captured by the investor.

Investment. Given the solution to the trade stage, we can now solve the problem of both the intermediary and investor being faced with a production opportunity. They both have to decide for which set of production costs c they are willing to produce. Notice that once investors reach the trade stage, the investment cost is a sunk cost, thus it does not directly affect the way the surplus is divided by the trade meeting participants. It is also straightforward to argue that the value of producing an asset for either the intermediary or the investor is higher the lower the cost actually paid to produce. This implies that both agents should produce the asset when they draw low costs and potentially neither produces the asset for high costs. Let c_l denote the threshold such that the intermediary produces if she draws a cost c below c_l . Likewise, let c_h denote the threshold such that the investor produces if she draws a cost c below c_h . Characterizing these two thresholds involves striking a balance between the benefits of investing and not investing.

Let $U_l(i|c_h)$ denote the utility of producing the asset for the intermediary when the investor's production threshold is given by c_h . Likewise, let $U_l(ni|c_h)$ denote the utility of not producing the asset for the intermediary when the investor's production threshold is given by c_h . A similar logic applies for describing $U_h(i|c_l)$ and $U_h(ni|c_l)$.

When the intermediary does not produce, we trivially obtain $U_l(ni|c_h) = 0$. When the intermediary produces at cost c , we get that

$$\begin{aligned} U_l(i|c_h) &= G(c_h)\nu_l + [1 - G(c_h)]p - c = \nu_h - G(c_h)(\nu_h - \nu_l) - \\ &\quad [1 - G(c_h)]\theta(\nu_h - \nu_l) - c , \end{aligned}$$

where the last expression follows by using equation (2). With probability $G(c_h)$, the investor produces an asset, and thus the intermediary keeps the asset for herself, providing her ν_l utiles. With probability $[1 - G(c_h)]$, the investor does not produce, which implies that trade will occur at the trading stage, and the intermediary obtains p .

The threshold c_l is such that if $c = c_l$, we get $U_l(ni|c_h) = U_l(i|c_h)$. Using the expressions above and solving for c_l provides

$$c_l = \nu_h - G(c_h)(\nu_h - \nu_l) - [1 - G(c_h)]\theta(\nu_h - \nu_l) . \quad (3)$$

Now we solve the problem of an investor. If the investor produces the asset, she gains the asset and does not trade later on. Thus, $U_h(i|c_l) = \nu_h - c$. If she does not produce the asset, her utility is given by

$$U_h(ni|c_l) = G(c_l)[\nu_h - p] = G(c_l)\theta[\nu_h - \nu_l] ,$$

where the last expression follows by using equation (2). The investor only gets utility from not investing when the intermediary invests, which occurs with probability $G(c_l)$. In this event, her gains from trade are $\nu_h - p$, or a fraction θ of the total trade surplus $\nu_h - \nu_l$. As before, the threshold c_h makes the investor indifferent. Then, we have that

$$c_h = \nu_h - G(c_l)\theta[\nu_h - \nu_l] . \quad (4)$$

From equations (3) and (4), we can compute the following expressions,

$$\frac{\partial c_l}{\partial c_h} = -g(c_h)(1 - \theta)(\nu_h - \nu_l) < 0 \quad , \quad \frac{\partial c_h}{\partial c_l} = -g(c_l)\theta(\nu_h - \nu_l) < 0 .$$

The first expression implies that if the intermediary expects the investor to be more likely to produce, then she will be less willing to produce herself. This occurs because the only reason the intermediary produces an asset is to attempt to sell the asset to the investor. If the probability of selling the asset falls (because the investor is more likely to hold an asset), there are lower incentives for the intermediary to produce. Likewise, because of the same logic, the second expression implies that the investor is less likely to produce if the intermediary is more likely to do it. That is, for both agents, their actions are strategic substitutes.

Equilibrium. An equilibrium is a pair of thresholds c_l^* and c_h^* that solve equations (3) and (4). Solving these two thresholds provides

$$c_l^* = \frac{[1 - \nu_h]\nu_h(1 - \theta)}{1 - \theta(1 - \theta)\nu_h^2} \in [0, \nu_h) \quad , \quad c_h^* = \frac{[1 - \theta(1 - \theta)\nu_h]\nu_h}{1 - \theta(1 - \theta)\nu_h^2} \in [0, \nu_h] .$$

Obviously, the fact that we were able to obtain closed-form expressions for both thresholds implies that the equilibrium exists and that it is unique. Moreover, the fact that the thresholds are interior and below ν_h follows from the fact that $\nu_h < 1$ and $\theta < 1$. The thresholds are not above ν_h , as there is always a profitable deviation that increases profits. For example, if $c_h^* > \nu_h$, the investor could trivially increase her expected utility by setting $c_l^* = \nu_h$. The fact that the threshold c_l^*

is strictly below ν_h follows from the fact that $c_l^* < c_h^*$. This follows, as there is a positive probability that the intermediary cannot sell the asset to the investor; because the intermediary values the asset at $\nu_l < \nu_h$, it follows that $c_l^* < c_h^*$.

2. EFFICIENCY

The constrained-efficient allocation is a set of thresholds c_l^e and c_h^e that solves the following problem,

$$\begin{aligned} \max_{c_l^e, c_h^e} \int_{\underline{c}}^{c_l^e} \int_{\underline{c}}^{c_h^e} [\nu_l + \nu_h - \tilde{c}_l - \tilde{c}_h] dG(\tilde{c}_h) dG(\tilde{c}_l) \\ + [1 - G(c_l^e)] \int_{\underline{c}}^{c_h^e} [\nu_h - \tilde{c}_h] dG(\tilde{c}_h) + [1 - G(c_h^e)] \int_{\underline{c}}^{c_l^e} [\nu_h - \tilde{c}_l] dG(\tilde{c}_l) , \end{aligned}$$

where the problem already assumes that if the intermediary produces an asset and the investor does not, the efficient outcome is for the intermediary to transfer the asset to the investor. There are three terms in the welfare expression. The first term accounts for the case where both agents draw a cost below their respective production thresholds and produce. Because both produce, both derive utility at the end of the period from holding the asset and both pay the production cost. The second term accounts for the case where the intermediary does not produce and the investor does produce. In this case, there is no trade and the investor keeps the asset she produces, obtaining utility ν_h minus the production cost incurred. The third term accounts for the case where only the intermediary produces. In this case, she pays the production cost and transfers the asset to the investor, who values the asset at ν_h .

The first-order condition with respect to c_l^e is given by

$$g(c_l^e) \left\{ \int_{\underline{c}}^{c_h^e} [\nu_l + \nu_h - c_l^e - \tilde{c}_h] dG(\tilde{c}_h) - \int_{\underline{c}}^{c_h^e} [\nu_h - \tilde{c}_h] dG(\tilde{c}_h) + [1 - G(c_h^e)](\nu_h - c_l^e) \right\} = 0 .$$

Operating with this expression provides

$$c_l^e = \nu_h - G(c_h^e)(\nu_h - \nu_l) . \quad (5)$$

Likewise, we can obtain the first-order condition with respect to c_h^e and operate to obtain

$$c_h^e = \nu_h - G(c_l^e)(\nu_h - \nu_l) . \quad (6)$$

Then, it is straightforward to obtain expressions for the thresholds,

$$c_l^e = c_h^e = \frac{\nu_h}{1 + \nu_h} . \quad (7)$$

Having solved for the constrained-efficient allocation, we are now ready to discuss the efficiency of the equilibrium. We do so in the next proposition.

Proposition 1 *For any $\theta \in [0,1]$, the equilibrium, characterized by $\{c_l^*, c_h^*\}$, is not constrained-efficient.*

This powerful result follows from the fact that there is no way of choosing the bargaining power parameter θ so as to satisfy $c_l^* = c_l^e$ and $c_h^* = c_h^e$. One way to prove this result is the following. Notice that in the constrained-efficient allocation we have that the thresholds satisfy $c_l^e = c_h^e$. Then, if the equilibrium were to be efficient, a necessary condition is that $c_l^* = c_h^*$. However,

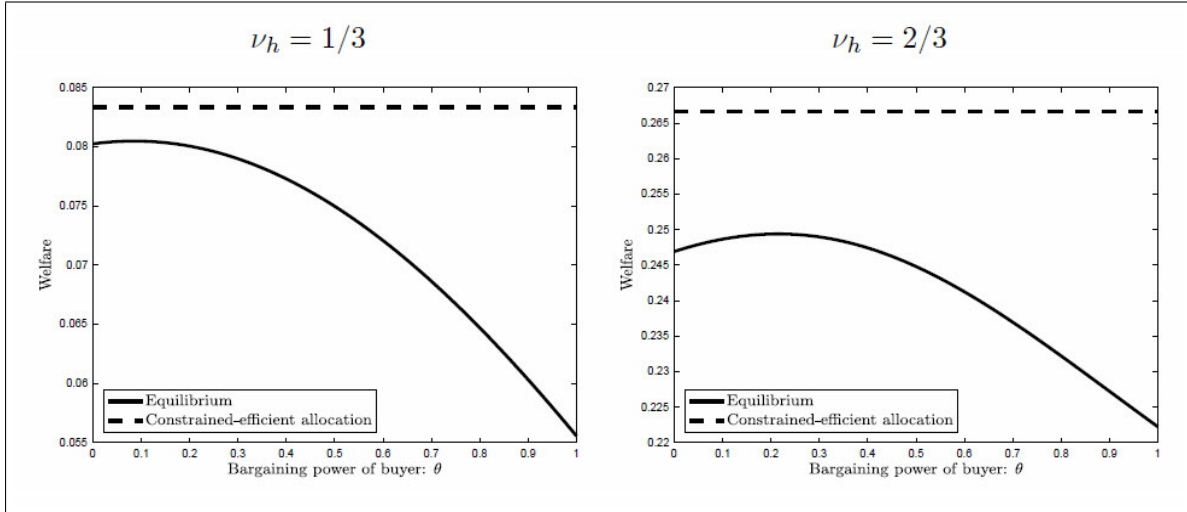
$$\frac{c_l^*}{c_h^*} = \frac{1 - \theta - \nu_h(1 - \theta)}{1 - \theta(1 - \theta)\nu_h} = 1 - \frac{\nu_h(1 - \theta) + \theta(1 - (1 - \theta)\nu_h)}{1 - \theta(1 - \theta)\nu_h} < 1,$$

so that $c_l^* < c_h^*$.

There is a key difference between the social values presented in equations (5) and (6) and private values presented in equations (3) and (4). When the planner evaluates the social value of allocating an asset to the intermediary, the planner takes into account that the intermediary will, with some probability, transfer the asset to the investor. In doing so, the planner takes into account the entire surplus generated by transferring an asset from the intermediary to the investor, or the entire surplus from trade. When the intermediary evaluates her reservation value, however, she takes into account only a fraction of the surplus. The reason is because she faces a hold-up problem—her decision to invest in producing an asset occurs before meeting with the buyer for the asset, i.e., the investor. The intermediary gives up c utiles (the cost of production) and gains the option of selling the asset to the investor. As usual in hold-up problems, the only way that the equations for the social value and reservation value of the intermediary coincide is if the intermediary has all the bargaining power when selling the asset to the investor. That is, $\theta = 0$. This is clear when we compare the equations describing the social and private values for the intermediary.

In an analogous way to what happens with the intermediary, when the planner evaluates the social value of the production by the investor, the planner takes into account that the investor will, with some probability, obtain the asset from the intermediary. In doing so, the planner takes into account the entire loss of surplus generated by passing an asset from the intermediary to the investor. When the investor evaluates her reservation value, she takes into account only a fraction θ , her bargaining power, of this surplus. The investor also faces a hold-up problem—her decision not to invest in producing an asset occurs before

Figure 1 Welfare in the Equilibrium Allocation and in the Constrained-Efficient Allocation



meeting with the intermediary. That is, in order to gain the option of buying the asset from the intermediary, she gives up the option of producing the asset herself, and this option is gone once she meets with the intermediary. The only way that the equations for the social value and reservation value of an investor coincide is if the investor has all the bargaining power when buying an asset. That is, $\theta = 1$.

Fixing the two sides of the hold-up problem is not possible, as this would require that both intermediary and investor have all the surplus generated by a trade. Both agents will never value the gains from trade in the same way the planner does, and as a result, the outcome of a decentralized equilibrium cannot replicate the planner's solution—no matter how investors bargain over gains from trade. Figure 1 presents the welfare of both equilibrium and constrained-efficient allocations for two different values of ν_h as we vary θ . As discussed, because there is no θ for which the constrained-efficient allocation can be decentralized, the equilibrium welfare lies below the constrained-efficient welfare for all $\theta \in [0, 1]$.

3. ASSET DISTRIBUTION

Through altering production incentives, the inefficiency manifests in the asset distribution of the economy. In this section, we aim to provide

some insights into how this happens. To this end, let s denote the asset level in the economy. There are three possible asset levels: 0, 1, or 2. There will be zero assets if both agents do not produce, there will be one asset if only one of them produces, and there will be two assets if both produce. To compute the expected asset level, we first need to compute the probability of each of the three potential outcomes. We have that

$$\Pr(s = 1) = (1 - G(c_l))G(c_h) + (1 - G(c_h))G(c_l) = c_l + c_h - 2c_l c_h, \quad \Pr(s = 2) = G(c_l)G(c_h) = c_l c_h,$$

and where $\Pr(s = 0) = 1 - \Pr(s = 1) - \Pr(s = 2)$. Then, the expected asset level is given by $E(s) = \Pr(s = 1) + 2\Pr(s = 2) = c_l + c_h$, which provides the following expressions for the equilibrium and efficient allocation,

$$E(s^*) = \nu_h \frac{2 - \theta - \nu_h(1 - \theta)(1 + \theta)}{1 - \theta(1 - \theta)v_h^2} > 0, \quad E(s^e) = 2 \frac{\nu_h}{1 + \nu_h} \in (0, 1).$$

We can combine these two expressions to obtain a relationship between the asset level in both cases:

$$E(s^*) = E(s^e) \frac{[2 - \theta - \nu_h(1 - \theta)(1 + \theta)](1 + \nu_h)}{2(1 - \theta(1 - \theta)v_h^2)}, \quad (8)$$

so that the equilibrium has a higher (lower) asset level than the efficient allocation if the second term on the right-hand side of the expression is larger (smaller) than one. Interestingly, this term can be above or below one, depending on the value for θ . In particular, when $\theta = 0$, we get that it is larger than one, and when $\theta = 1$, we get that it is smaller than one. This implies the following result.

Proposition 2 *The equilibrium allocation exhibits a lower asset level than in the constrained-efficient allocation when $\theta = 1$, and it exhibits a higher asset level than in the constrained-efficient allocation when $\theta = 0$.*

In equilibrium, when $\theta = 1$, all the gains from trade go to the investor. Thus, the intermediary has no incentives to produce the asset. However, the planner would like the intermediary to produce the asset with some probability, as this increases the chances that the investor ends up holding the asset. As a result, the asset supply is inefficiently low when θ is high. When $\theta = 0$, the intermediary has maximum incentive to produce the asset (as she gets all the gains from trade) but so does the investor, as she would get no gains from trading with the intermediary. As a result, the asset supply is inefficiently high when θ is low. Figure 2 presents the asset level in the equilibrium

and constrained-efficient allocations for two levels of ν_h . As seen in the figure, consistent with the discussion about $\theta = 0$ and $\theta = 1$, for low values of the bargaining parameter θ , the equilibrium exhibits overproduction, while for high values of θ , it exhibits underproduction.

The figure also shows that there is a value for the bargaining power such that the equilibrium asset supply equates the asset supply in the constrained-efficient allocation. However, given our impossibility result in Proposition 1, we know that the allocation is inefficient even when the equilibrium asset supply matches the efficient asset supply. What happens is the distribution of assets across the two agents is inefficient, in the sense that even though we may find a θ such that $E(s^*) = E(s^e)$, at this θ it must be the case that the equalities required for efficiency, i.e., $c_l^* = c_l^e$ and $c_h^* = c_h^e$, are not satisfied. In fact, given that $c_l^* < c_l^e$ and $c_l^e/c_h^e = 1$, it is the case that, for a given asset level, the way assets are distributed across the two agents is never the same as in the efficient allocation. This observation allows us to provide an alternative interpretation of the inefficiency result. The double hold-up problem affects efficiency by distorting both the level of the asset supply and the way assets are distributed across agents. There is always a way of choosing θ in order to fix the first distortion: from equation (8), and by continuity argument, it is clear that there always exists a way to choose θ such that $E(s^*) = E(s^e)$. However, given $c_l^*/c_h^* \neq c_l^e/c_h^e$ for all θ , there is no way of correcting the second distortion.

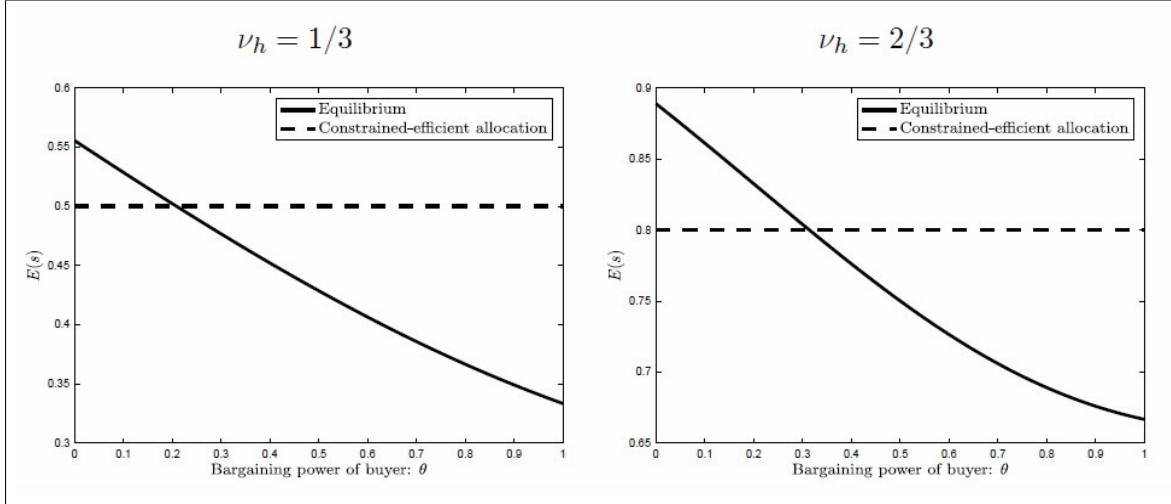
4. AN EXAMPLE WHERE THE EQUILIBRIUM IS CONSTRAINED-EFFICIENT

Although the model we developed here has no search frictions, in the sense that at the production stage agents know who they will meet at the trading stage, there is a fundamental friction in the model that stems from the fact that both parties in trade, when faced with the opportunity to produce, are unaware if their trade counterparty will be holding an asset or not when they meet at the trade stage. This friction is important for understanding the inefficiency result presented in Proposition 1. To study this in some detail, we allow the upper bound of the production cost distribution \bar{c} to be in the set $[\nu_h, 1]$. Recall that, before, we had assumed $\bar{c} = 1$. Notice that the inefficiency results discussed above hold for $\bar{c} > \nu_h$. Thus, think of the case $\bar{c} = \nu_h$ as the limiting case when $\bar{c} \downarrow \nu_h$.

When $\bar{c}_h \in [\nu_h, 1]$, the equilibrium thresholds, equations(3) and (4), reduce to the following expressions,

$$c_l^* = \nu_h(1 - \theta) \left(1 - \frac{c_h^*}{\bar{c}} \right) \quad , \quad c_h^* = \nu_h \left(1 - \frac{c_l^*}{\bar{c}} \theta \right) .$$

Figure 2 Asset Level in the Equilibrium Allocation and in the Constrained-Efficient Allocation



Likewise, the thresholds in the constrained-efficient allocation, equations (5) and (6), reduce to

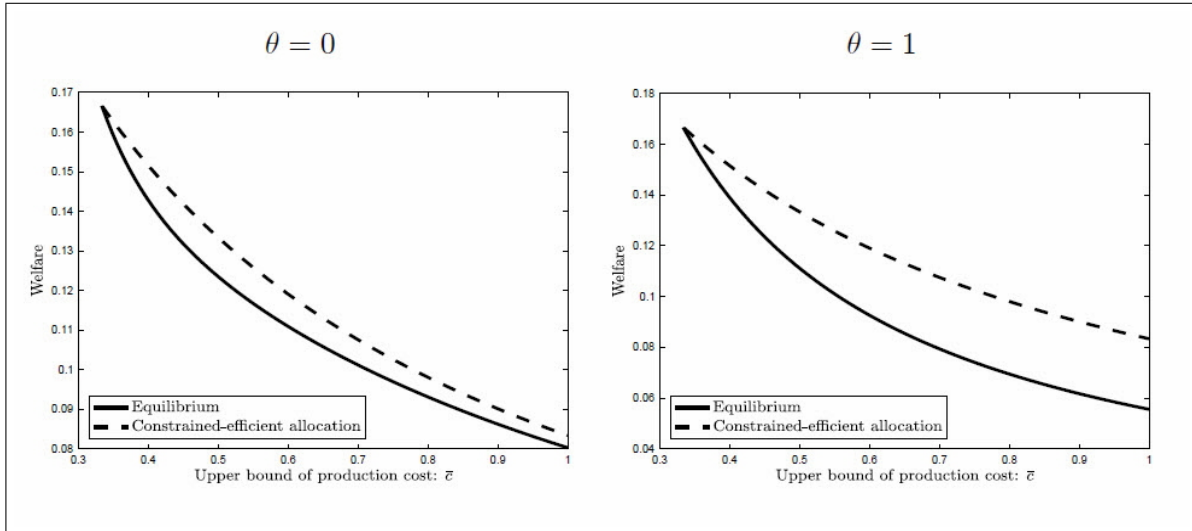
$$c_l^e = \nu_h \left(1 - \frac{c_h^e}{\bar{c}} \right) , \quad c_h^e = \nu_h \left(1 - \frac{c_l^e}{\bar{c}} \right) .$$

Suppose $\bar{c} = \nu_h$. In the constrained-efficient allocation, we obtain $c_l^e + c_h^e = \nu_h$. Because the production cost distribution is uniform, the density is the same at every production cost. As a result, the efficient allocation cares about maximizing the production probability, which equals ν_h , but it is silent on how production is divided between the intermediary and the investor. Notice that the efficient allocation can be attained with trade, for example, if only the intermediary produces ($c_l^e = \nu_h$ and $c_h^e = 0$), with cases where trade sometimes happens (both agents having interior production thresholds), or cases with no trade (which occurs when only the investor produces).

In particular, the planner can attain the constrained-efficient allocation by setting $c_l^e = 0$ and $c_h^e = \nu_h$: the intermediary does not produce, and all production is carried over by the investor. Now we turn to the equilibrium thresholds. Combining c_l^* and c_h^* provides an expression for c_h^* ,

$$c_h^* = \frac{1 - \theta(1 - \theta)}{\frac{1}{\nu_h} - \frac{\theta(1 - \theta)}{\nu_h}} = \nu_h ,$$

Figure 3 Welfare and the Upper Bound of Production Cost Distribution $\nu_h = 1/3$



where we already used that $\bar{c} = \nu_h$. Also, we immediately obtain that $c_l^* = 0$. In other words, when $\bar{c} = \nu_h$, the equilibrium is constrained-efficient for any $\theta \in [0, 1]$. This occurs because by reducing \bar{c} up to ν_h we are taking to zero the probability that the investor does not produce. This implies that the intermediary can forecast perfectly that the investor will arrive to the trade meeting holding an asset. Likewise, because $c_l^* = 0$, the investor can forecast perfectly that the intermediary will arrive to the trade meeting not holding an asset. Thus, the intermediary does not produce as she knows she cannot offload the asset to the investor, and the investor produces as she knows she cannot acquire the asset from the intermediary. Figure 3 presents this result for this simple example. It shows that as \bar{c} approaches $\nu_h = 1/3$, the equilibrium welfare approaches the welfare of the constrained-efficient allocation. It shows this in the left panel for the $\theta = 0$ case and in the right panel for the $\theta = 1$ case.

5. CONCLUSION

In this paper, we show that there is a generic source of inefficiency in OTC markets stemming from the fact that agents must make production choices (or, in general, trade decisions) ex ante without knowing

the choices of their trade counterparties ex post. Due to the lack of a large mass of agents, individual actions have strong effects on other agents' decisions. This implies that the equilibrium is not constrained-efficient, independent of the way agents split the surplus among them.

We then show that the inefficiency result vanishes in an example where the equilibrium provides more information to market participants at the production stage, thus improving their forecasting power regarding the actions of other players in the market. In the example we study, when the investor always invests, the intermediary never invests, so agents know with certainty if their trade counterparties will be holding an asset or not when they meet. We show that this attains efficiency.

Certainly, the model is a simple one. However, the inefficiency result seems robust. Given the growing interest in understanding trading in OTC markets, we believe that understanding the welfare properties of models of OTC trading is of paramount importance and calls for further research.

REFERENCES

- Babus, Ana, and Péter Kondor. 2013. "Trading and Information Diffusion in OTC Markets." Centre for Economic Policy Research Discussion Papers 9271 (January).
- Bethune, Zachary, Bruno Sultanum, and Nicholas Trachter. 2017. "Asset Issuance in Over-the-Counter Markets." Federal Reserve Bank of Richmond Working Paper 17-13 (October).
- Duffie, Darrell, Nicolae Garleanu, and Lasse Heje Pedersen. 2005. "Over-the-Counter Markets." *Econometrica* 73 (November): 1815–47.
- Duffie, Darrell, Nicolae Garleanu, and Lasse Heje Pedersen. 2007. "Valuation in Over-the-Counter Markets." *Review of Financial Studies* 20 (November): 1865–1900.
- Farboodi, Maryam, Gregor Jarosch, and Guido Menzio. 2016. "Intermediation as Rent Extraction." Penn Institute for Economic Research Working Paper 16-026 (December).
- Farboodi, Maryam, Gregor Jarosch, and Robert Shimer. 2017. "Meeting Technologies in Decentralized Asset Markets." Working Paper (February).
- Gofman, Michael. 2014. "A Network-Based Analysis of Over-the-Counter Markets." Working Paper (December).
- Hosios, Arthur J. 1990. "On the Efficiency of Matching and Related Models of Search Unemployment." *Review of Economic Studies* 57 (April): 279–98.
- Hugonnier, Julien, Benjamin Lester, and Pierre-Olivier Weill. 2014. "Heterogeneity in Decentralized Asset Markets." Working Paper 20746. Cambridge, Mass.: National Bureau of Economic Research (December).
- Lagos, Ricardo, and Guillaume Rocheteau. March 2009. "Liquidity in Asset Markets with Search Frictions." *Econometrica* 77 (March): 403–26.
- Lagos, Ricardo, and Randall Wright. 2005. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy* 113 (June): 463–84.