A Dixit-Stiglitz general equilibrium model with oligopolistic markets: Enough is Enough

> John P. Conley Vanderbilt University

Robert Driskill Vanderbilt University

Southern Economic Association Meeting

November 2015

Introduction

The Dixit-Stiglitz (1977) model of Chamberlainian monopolistic competition model and has had a enormous impact on research in Industrial Organization, the Economics of Geography, Monetary and Real Business Cycle, Growth Theory, and International Trade.

Melitz (2003) translated this model into a measure space. Melitz's model has proven to be both tractable and flexible and has itself had a large impact on the international trade literature (cited more than 3000 times according to Google scholar).

Introduction

We puzzled about what this paper was trying to say.

Our starting point was really a purely mathematical concern about how Melitz's approach corresponds to the underlying economy it seeks to model.

As anyone knows who has worked in measure spaces, something that is mathematically correct may not be economically sensible.

To be sensible, the continuum outcome should be the limit of a large finite economy.

Introduction

A classic example of how the mathematical features of the continuum can yield artificial results can be be seen the marriage problem.

Suppose there is an interval [0,1] of girls, and [0,2] of boys. Note that if I match girl number g to boy number 2g, each and every agent has a partner.

Clearly, this is nonsense. If the large finite economy has twice as many boys as girls, half the boys must remain unmatched. The continuum model, in other words, has not reflected the economic fact that girls are scarce. Thus, there is very little predictive value in continuum model in this case.

(See Kaneko and Wooders (1986. 1989) on "Measure Compatibility" for a solution.)

What about Melitz? His approach is to use a representative consumer whose preferences are given by a C.E.S. utility function over a continuum of goods indexed by ω

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\rho} d\omega \right]^{1/\rho}$$

where the measure of the set Ω represents the mass of available goods.

Thus: there are an uncountable infinity of goods, each produced by a single firm.

Then $P = \left[\int_{\omega \in \Omega} p(\omega)^{1\sigma} d\omega\right]^{1/1-\sigma}$ is the price of the "aggregate good" Q where $\sigma = 1/(1-\rho) > 1$ and:

$$q(\omega) = \left[Q\frac{p(\omega)}{P}^{-\sigma}\right]$$

Note that this means that positive levels of each good are consumed by the representative agent

The question then is how to understand what is meant by this.

A: Perhaps the economy consists of an infinity of homogeneous agents with utility functions exactly like the representative agent above. What are the problems with this?:

1. This is resource infeasible: Suppose good .3 is a banana, and each agent consumes one banana. If I integrate to find the number of bananas consumed, a measure one of bananas must be produced. If it costs one unit of labor per banana, a measure one of labor is required. Thus, the measure of bananas is of the same order of magnitude as the total supply of labor. Then since all agents consume a finite amount of an infinity of goods, an infinite measure of labor is required. This is a slight violation of the resource constraint. (not to mention, agents' stomachs would explode!)

2. Another way to see this is that ask: what it would cost an individual to buy a finite quantity of an infinity of goods? If the prices are also finite, then the cost is infinity, a slight violation of the budget constraint.

3. A way out of might be for the consumption level of good to go to zero as the number of goods increases (the limit of a finite model). (The Melitz model does not take this approach, and so his continuum model apparently does not reflect the limit of a large economy such as this.) However, we would have to resort to a non-standard analytic approach to reflect that demand for each good would literally be infinitesimal, but when multiplied by prices and integrated, expenditure equals income. Messy, perhaps possible, but not what we see in Melitz and of doubtful economic relevance anyway.

B: Perhaps the utility function is really just a reflection of aggregate behavior (not average), and so the resulting demand is market demand.

4. In this case, feasibility is restored. Hooray!

5. This utility function, however, cannot be representative of a homogeneous and infinite set of consumers. If so, we would face problem 3 above at the individual level. Thus, to be sensible, the utility function must represent the aggregate behavior of an infinite set of heterogeneous consumers. This means there must be some micro-foundation with heterogeneous agents behind the Melitz approach. The correct interpretation of Melitz's model and results therefore depend upon this exactly what the micro-foundation is.

Our Objectives

The central question addressed in this paper is: What is the right micro-foundation? Our objective is therefore to:

I. Establish a sensible micro-foundation for a continuum oligopolistic competition model.

II.Explore a model based on these micro-foundations to see if the economic policy conclusions are in any significant way different from Melitz type models, especially for international trade.

Our Objectives

First we make a key observation: Individual agents (even in a large finite economy) consume at most a finite (in fact, a strictly bounded) set of goods. Thus, agents can be seen as first identifing a set set of goods, and second choosing how much of each to consume,

For example, how many of the goods offered at Amazon have your consumed? One reason that you only purchase a tiny fraction of Amazon's offerings might be transactions costs. Even if I had infinite money and stomach capacity, I would not have the time to complete an infinity of transactions.

A second reason might be non-convexities driven sometimes by indivisibilities. Half a Camery plus half an Accord won't take as far or as fast as whole Hyundai.

We consider an economy with an uncountably infinite set (endowed with Lebesgue measure) of heterogeneous consumers in the economy and denote agents by the index $i \in [0, I] \equiv \mathcal{I}.$

Agents are each endowed a strictly bounded quantity the non-produced good: $\omega_i \leq \overline{\omega} > 0$. Let $\Omega(i)$ be a measurable function that gives the endowment for the economy.

Agent also consume good produced by firms. Produced goods each have a characteristic g drawn from metric space (G, d) where G is a convex and compact set.

The cost of producing a good g, is given by an affine function:

$$C^g(Y^g) = F^g + V^g Y^g,$$

where F^g and V^g are independently drawn the intervals $[\underline{F}, \overline{F}]$ and $[\underline{V}, \overline{V}]$ under frequency distributions f^F and f^V , respectively where $\underline{F}, \underline{V} > 0$.

The planner undertakes R&D projects at a cost of D each.

Researching a product characteristic, g, allows him to produce the corresponding good under a cost function outlined above.

The planner may therefore choose to research a product characteristic multiple times in hopes of finding how to produce it more cheaply.

We assume that agents can always consume the non-produced good, but deciding to consume a positive quantity of any produced good, and thus, adding it to ones' utility function, is costly.

Thus, if an agent *i* decides to add a produced good *g* to his demand set, \mathcal{N}_i , he must pay a cost of t > 0 of the non-produced good to cover the transactions costs.

The size of the demand set $|\mathcal{N}_i| = N_i$ is therefore endogenously determined and may differ across agents.

Given this, the utility function of the agent i is given by the quasi-linear function:

$$U_i(x_i, y_i) = y_i + h_i(\{x_i^g\}_{g \in \mathcal{N}_i}) - tN_i$$

We make two key assumptions on agents and goods.

Assumption A. Agents in a close neighborhood of each other are similar: For all $\epsilon_{\mathcal{I}} > 0$ there exists $\delta_{\mathcal{I}} > 0$ such that for all $i, j \in \mathcal{I}$ such that $|i - j| < \delta_{\mathcal{I}}$ then

- (a) $|\omega_i \omega_j| < \epsilon_{\mathcal{I}}$
- (b) for any demand set $\mathcal{N}_i \subset G$ and consumption bundle (x_i, y_i) permissible under \mathcal{N}_i , $|U_i(x_i, y_i) U_j(x_i, y_i)| \leq \epsilon_{\mathcal{I}}$

This says that if agents are close together as measured by the their index then they have similar endowments and get similar utility levels from the same consumption bundles.

Assumption B. Goods in a close neighborhood of each other are close substitutes: For all $\epsilon_G > 0$ there exists $\delta_G > 0$ such that for all $i \in \mathcal{I}$ and all integers N: (a) for all $\overline{\mathcal{N}_i} = (\bar{g}_1, \dots, \bar{g}_N), \widehat{\mathcal{N}_i} = (\hat{g}_1, \dots, \hat{g}_N)$ such at for $n = 1, \dots, N, d(\bar{g}_n, \hat{g}_n) < \delta_G$ and $x_i^{\bar{g}_n} = x_i^{\hat{g}_n}$ then $|U_i(\bar{x}_i, y_i) - U_i(\hat{x}_i, y_i)| < \epsilon_G$. (b) for all $\overline{\mathcal{N}_i} = (\bar{g}_1, \dots, \bar{g}_N)$, if for some $\bar{g}_n \in \overline{\mathcal{N}_i}, \widehat{\mathcal{N}_i} = \overline{\mathcal{N}_i} \setminus \bar{g}_n$ and for some $\bar{g}_m \in \overline{\mathcal{N}_i}$ $d(\bar{g}_n, \bar{g}_m) < \delta_G$ it is the case that $x_i^{\hat{g}_m} = x_i^{\bar{g}_m} + x_i^{\bar{g}_n}$ while for all $k \neq n, m, x_i^{\bar{g}_k} = x_i^{\hat{g}_k}$ then $|U_i(\bar{x}_i, y_i) - U_i(\hat{x}_i, y_i)| < \epsilon_G$.

This says that if we replace each good in a demand set with an identical levels of similar goods as measured by d, then utility levels change very little. Also, utility change is small if we consolidate consumption of two similar goods on one of them.

Assumption C. Continuity and diminishing utility of produced goods: For all $i \in \mathcal{I}$ h_i is continuous and there exists an upper bound \overline{B} such that for any consumption vector (x, y) that include some good g for which $x^g > \overline{B}$, it holds that: $\partial h(x)/\partial x^g < \underline{V}$

Note that convexity and monotonicity are not needed. We only require that the marginal willingness to pay for any produced good falls below the minimal marginal cost of production at some high consumption level. Without this, it might be possible for an agent to get infinite utility by allowing the consumption level of the endowment good to be infinitely negative in order to pay or infinitely high levels of produced good in which his is never satiated.

Finally, we assume there is at least a positive probability any given pair of cost parameters will be draw for any particular g.

Assumption D. Full support in cost space: The frequency distributions f^F and f^V over $[\underline{F}, \overline{F}]$ and $[\underline{V}, \overline{V}]$ respectively are strictly positive.

A feasible plan is $(\mathcal{N}^D, \mathcal{N}, X, Y, C)$ Where $\mathcal{N}^D \subset G$ is the choice of research projects to undertake, $\mathcal{N} \subset \mathcal{N}^D$ is the choice of goods to actually produce, $Y : \mathcal{I} \to \Re$ and $X : \mathcal{I} \to \Re^N_+$ are measurable functions describing the allocation of goods to agents, and $C : \mathcal{I} \times \mathcal{N} \to \{0,1\}^N$ is a measurable function that describes whether or not agent *i* has added good $g \in \mathcal{N}$ to his demand set, such that:

$$\int Y(i)di + \sum_{g \in \mathcal{N}} \left(F^g + V^g \int X^g(i)di \right) + \sum_{g \in \mathcal{N}} t \times C(g,i)di - D \times |\mathcal{N}^D| \le \int \Omega(i)di$$

In the following, it will be useful to define the notion of a minimal δ -grid in the goods space G. Recall that G is bounded. Let \mathcal{N} be a finite subset of G. then \mathcal{N} is a δ -grid of G if

$$\max_{g \in \mathcal{N}} \min_{g \in \mathcal{N} \atop \bar{g} \neq g} d(g, \bar{g}) \le \delta$$

Let \mathbf{N}^{δ} be the set of all δ -grids of G. Then the \mathcal{N} is a minimal δ -grid of G if no $\overline{\mathcal{N}} \in \mathbf{N}^{\delta}$ has a smaller number of elements. Note this is well defined since the set of elements are integers and are bounded below.

Intuitively: a δ -grid of G is the smallest collection of goods that assures that no good is more than δ from at least one other good.

The question we want to address is what are the potential benefits from product diversity and therefore trade. Thus, we consider this from the planner's standpoint instead of solving for equilibrium.

If the equilibrium, whatever it is, if Pareto optimal, then it achieves the socially optimal welfare (given quasi-linearity). If it is not, then the planner's solution is an upper bound of social welfare.

Our strategy here is not exactly to solve the planner's problem.

Instead, we construct a series of well-defined and feasible plans. Since these plans are feasible, each one provides a lower bound on the social welfare one would obtain in any socially optimal plan. We show:

We show that the welfare obtained in these "guarantee point" plans converges to maximum possible social welfare as the population gets large.

This implies that the social payoff to increasing the number of goods produced diminishes to zero in the limit. This in turn implies that trade between any two economies diminishes to zero as the economies both get large.

Let's begin by considering how well an individual consumer i could conceivable do in the best of all possible worlds. Thus, we consider the following idealized consumer's problem: Suppose that an agent could choose any demand set $\mathcal{N}_i \subset G$ and could purchase as much of each good as he wished at price \underline{V} per unit. Formally the idealized consumer consumer problem is:

$$\max_{\mathcal{N}_i \subset G} \max_{\{x_i^g\}_{g \in \mathcal{N}_i}} h_i(\{x_i^g\}_{g \in \mathcal{N}_i}) - tN_i - \underline{V}^g x^g.$$

We will call a solution an *idealized consumer demand* and denote it as: \mathcal{N}_i^* and x_i^* respectively. We now show two things about the about this solution.

Lemma 1. There exists an integer \overline{N} such that for all $i \in \mathcal{I}$ and for any solution to the the agent's idealized consumer problem, $N_i < \overline{N}$

Intuition: Given that goods in a close neighborhood are close substitutes, the gain from splitting consumption over two similar goods gets small. If the goods are close enough, the gain must be smaller than t the cost of adding goods to the demand set. Since the innovation space G is compact, if agents add an unbounded number of goods to their demand sets, eventually some of these goods must be extremely close to each other. It would then be optimal to consolidate consumption on a smaller set of goods.

Lemma 2. There is an upper bound \overline{U} such that for all $i \in \mathcal{I}$ utility that any agent can receive at any solution to the the agent's idealized consumer problem is less than \overline{U} .

Intuition: Assumption C says that utility is continuous and that agents are eventually "satiated" in produced goods. Thus, agents will choose at most a finite amount of a finite set of goods. Since utility is also continuous over the compact innovation space G, the consumer's problems is to maximize a continuous function over a compact set. This has a finite solution.

What the two Lemmas above show is that there is a theoretical upper bound on the utility that any agent in the economy can obtain and that this involves the agent consuming at most a strictly bounded set of goods. Of course, it may not be feasible to have sufficient product diversity to allow all agents to achieve this maximum, and even if it was, the resulting allocation might or might not be measurable. The next Lemma will show that despite this, there is a plan that allows each agent to get arbitrarily close in expectation to his theoretical maximum utility which is feasible for a large enough economy.

A Generalized Model (Skip)

Thus, we consider how the planner can generate a feasible plan that approximates these idealized utility levels for each agent. Define a (ϵ, R) -plan as follows:

- 1. \mathcal{N}^D : Let $\overline{\mathcal{N}}$ be any minimal δ -grid of G where $\delta = \delta_G$ and δ_G is associated with $\epsilon_G = \epsilon$. Then the set of research and development projects undertaken, \mathcal{N}^D , is defined as to developing each project in $\overline{\mathcal{N}}$ a total of R separate times.
- 2. \mathcal{N} : Let $\{g_1, \ldots, g_R\}$ be the *R* separate tries at development in \mathcal{N}^D of any given product in $g \in \overline{\mathcal{N}}$. Then for each $g \in \overline{\mathcal{N}}$ construct \mathcal{N} choosing a $g_i \in \{g_1, \ldots, g_R\}$: with the lowest average realization of V^{g_i} and F^{g_i} : $\frac{V^{g_i} + F^{g_i}}{2}$.

A Generalized Model (Skip)

- 3. C: Divide the interval of agents into consecutive coalitions of agents each ϵ units long. For each of these intervals, take the middle agent. Choose any one of the idealized consumer demands for agent $i, \mathcal{N}_i^*, x_i^*$. For each $g \in \mathcal{N}_I^*$ We will say that good in $\bar{g} \in \mathcal{N}$ is in correspondence with good g if it is closest under d: $\hat{g} = \operatorname{argmin}_{g \in \mathcal{N}} d(g, \hat{g}).$ (In the zero probability event that two goods in \mathcal{N} are the same distance from g choose one of them randomly.) Given this, define \mathcal{N}_i to be the set of goods in \mathcal{N} that are in correspondence with the goods in \mathcal{N}_i^* . Having established the demand set for the middle agent of each interval, we now set The demand set of every agent j who shares the interval with agent i shares to be the same: $\mathcal{N}_j = \mathcal{N}_i^*$ Finally, C, the indicator function for which agents are consuming which produced goods is just constructed to reflect these demand sets over intervals consumers.
- 4. X: For each agent *i* in the middle of an interval we set x_i as follows. For every $\hat{g} \in \mathcal{N}$, set x_i^g to equal the sum consumption level of each of the goods $g \in \mathcal{N}$ that \hat{g} is in correspondence with. (again, for a fine grid, the correspondence will be one-to-one).

As above set the consumption levels of each agent j who shares the interval with i to be the same $x_j = x_i$. Finally, X, the produced good consumption mapping, is constructed to reflect these intervals of identical consumption levels.

A Generalized Model (Skip)

5. Y: For each agent i in the middle of an interval, set

$$y_i = \omega_i - \frac{RND + \sum_g F^g}{I} - \sum_{g \in \mathcal{N}_i} V^g x_i^g - tN_i.$$

That is, his endowment minus his average share of the total of the development and fixed production costs in the plan minus the variable cost of his produced goods consumption minus the attention cost of having a consumption set of size N_i . In addition for every agent j who shares the interval with agent i, let

$$y_j = \omega_j - \frac{RND + \sum_g F^g}{I} - \sum_{g \in \mathcal{N}_j} V^g x_j^g - tN_j \equiv \omega_j - \frac{RND + \sum_g F^g}{I} - \sum_{g \in \mathcal{N}_i} V^g x_i^g - tN_i.$$

Thus, the consumption of the non-produced goods for agents in a given interval differs only by the difference in their initial endowment of the good.

Intuitively, a (ϵ, R) -plan breaks the set of agents into a series of intervals each ϵ long.

The planner chooses a δ -grid (which gets finer as ϵ goes to zero) and researches each good g in this grid with a view to finding lower fixed and variable costs.

The planner then takes the agent in the middle of each interval and give him a demand set and consumption levels that are as close as possible, given the δ -grid, to his idealized demand

We deduct the variable costs of the middle agent's consumption and also an equal share of fixed and development costs from his endowment to find his consumption of y

Finally, we give all agents in a given interval the same allocation as the middle agent.

Lemma 3. Any (ϵ, R) -plan is feasible.

Intuition: The challenge is to show that the allocations are measurable functions. This basically comes from the fact the agents in a measurable interval all get the same allocation. Thus, the allocations are bounded step functions which in turn are measurable.

Lemma 4. For any $\gamma > 0$ there exists a (ϵ, R) -plan such that for a large enough population, each agent in \mathcal{I} receives within γ of the utility he would get at an idealized demand.

Intuition: Consider an agent *i* in the middle and the interval for some (ϵ, R) -plan. There are three things that might prevent him from getting his idealized utility:

- 1. It may be that he is forced to consume a set of goods from the δ -grid are slightly less desired than the goods in his idealized demand.
- 2. It may be that he pays $V^g > \underline{V}$ per unit of some goods he consumes.
- 3. It will be the case that his share of fixed and development costs is positive and thus he will receive this much less non-produced good, *ceteris paribus*

None of these bite in the limit.

Theorem 1. (Enough is enough) The per capita benefits of having greater product diversity diminishes to zero no matter how large the economy.

Proof: From Lemma 4, we know see that we can get to within any γ of the idealized utility level of every agent for a large enough economy while producing \overline{N} goods or fewer. Thus, no matter how much the economy continues to grow, at best, increasing the product set adds at most γ per capita utility. On the other hand, for any given economy of fixed size, increasing the number of goods eventually adds more to per capita development and fixed production costs than it does to utility. We conclude that for any economy, enough is enough.

Finally, consider two parallel economies and assume there are iceberg costs of $m \in (0, 1)$ such at only a fraction of m of exported goods arrive at their destination.

Theorem 2. It is impossible for a positive fraction of agents to consume any imports as both economies get arbitrarily large.

Intuition: The exporting country must sell goods at at least $\frac{V}{m}$ to overcome iceberg costs. If the a positive positive fraction of agents to consume any import and the importing country is large enough, and then it becomes economic to do enough development efforts to lower the price of domestically produced substitute goods below the costs of imports and share the costs of development and the fixed costs of production over the arbitrarily large group who consume these imports. Since these domestically produced goods are close substitutes, cheaper on the margin, and contributes arbitrarily little to average shares of fixed cost, they would crowd out the imports.

One implication is that in the micro-framework outlined above, large contries have little incentive to trade Such counries already have enough product variety provided domestically, and so would not be. eilling to pay iceberg costs for more product variety.

Clearly we do trade, however. Thus, we must ask: Why? The model suggests that reasons for trade are actually more classical:

1. Trade should take place when fixed costs are large compared to the size of the economy. Consider aircraft. Fixed costs of development and setting up production lines are so large, that we only have a very few producer of large commercial aircraft. Most counties import such aircraft. We have perhaps a hundred or so car manufactures in the world, but evidently, the five or ten producers of cars in even the largest country are not able to provide sufficient product diversity to foreclose welfare enhancing trade. Thus, our point is while a desire from product diversity does in fact motivate trade, this desire is ultimately bounded. Thus, it should only be economically relevant when fixed costs are sufficiently high in comparison to the market demand that only a relatively few firms can produce in the product category in equilibrium.

2. Trade should take place when there are differences in abilities. Many French cheese and wine producers are relatively small, and there are many thousands of wine and cheese produces in most large countries. However, we still trade wine and cheese with the French. In the model, this could be explained by the French having a monopoly over part of the innovation space. That is, for technological reasons, there is a region of G that can only be produced in France. If Americans have preferences that make it optimal for some goods g in this region to be added to their demands sets, they can only import these goods from France. Of course in reality, we do see these technological specializations. The model suggests we should expect to see trade even when there are many producers and fixed costs are low in these cases.

3. Trade could take place if certain countries are technological leaders. If one takes the view that (a) new innovations deprecate old ones or (b) the boundary of the innovation space G expands each year, then countries who happened to have lower development costs D would have a systematic advantage in bringing out new products. Since less technically apt countries would not innovate as much, there would be a flow new goods from the high tech countries to the low tech countries. In contrast, we assumed that all countries were equally able to produce new developments in our model.

4. Trade should take place when there are differences in endowments. China has cheap labor due to its factor endowments. As a result, the assumption that each country is drawing from the same fixed and variable cost distribution for new innovations is not empirically correct. Even though we have the technology to produce any of the manufactured goods that China sends us, China can produce these at a sufficiently lower price that it can overcome the iceberg costs. Resources may also provide literal monopoly or oligopoly positions in exports. Only a few countries can export diamonds or uranium. Formally, one would model this as certain countries having exclusive abilities to produce goods in certain regions of G.

Thus, the project of trying to build a case of trade based on desire for product variety when we have a continuum (or even a large number) of oligopolistic firms seems to be fundamentally misguided.

Deep Thoughts

What is the nature of innovation?

We have treated it as choosing product characteristics from a compact set. This means there is really a natural limit on the how well off I can be no matter how closely the produced products match my ideal goods (at least under the assumptions that close products are similar and that we have limits of the number of goods we can consume).

An alternative is a product quality ladder type innovation structure with goods getting potentially unboundedly good. However, this would suggest that as time goes on, we get happier and happier. People in the middle ages would also have been comparatively unhappy. This does not seem to agree with the psychometric evidence.

Deep Thoughts

Perhaps new innovations deprecate the value of existing products. When I see the new iPhone, I start to realize how lame the old one was.

But then, innovation is actually destructive. It wastes resources and in the end just replaces old goods with new goods that don't make us any happier.

Nevertheless, innovation will still take place in competitive markets. Thus, innovation is a kind of market failure.

Conclusion?

All research, including this work is socially inefficient and should stop immediately.

Conclusions for Trade

First: Trade is most likely to be relevant for goods produced with large fixed costs. Thus, it may simple be wrong headed to try to construct a model in which trade takes place with infinitesimal firms, or equivalently, with infinity many products.

Second: For goods which are produced on a small scale by (monopolistic) competitive industries, trade takes place on the basis of comparative advantage that is related to location, and not due to idiosyncratic differences in cost within a country.

Third: Even for a finite Dixit-Stiglitz type model, it is still necessary to bound the number of goods agents consume if one wishes to understand a large economy.