Welfare in models of trade with heterogeneous firms

Alexandre Janiak*
University of Chile

Abstract

I illustrate that the welfare improvement property of the Melitz model is due to the shape of the aggregate labor demand curve, which slopes upwards. By slightly changing some assumptions in the model, this curve may have a negative slope. In this case, increases in aggregate productivity result in a reduction in welfare. For example, this may occur when fixed costs are measured in units of aggregate output instead of labor.

JEL codes: F12, F16, J23.

Keywords: heterogeneous firms, international trade, aggregate labor demand curve, welfare.

*Contact details: Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Universidad de Chile, República 701, Santiago, Chile. Tel: +56 2 978 4912. E-mail: ajania@dii.uchile.cl
1 Introduction

The Melitz (2003) model has been extensively used in international economics to the point that it is now a cornerstone of the field. On top of adding a new dimension to economic modeling - heterogeneity in productivity - Melitz's model does it in a very clear manner. The model is a simple extension of Krugman (1980), where production is characterized by scale economies and heterogeneity arises because a firm's productivity is drawn from an exogenous distribution function. While many models in macroeconomics and consumption theory require the use of computers to solve for equilibrium, this model provides a close-form solution, which makes the debate in economics easier. In particular, it has become easier to study the impact of competition policies, such as deregulation or trade barriers, on aggregate productivity.

This paper is a comment on the welfare implications of the model. The main channel through which trade liberalization improves aggregate welfare is labor reallocation. As only the most productive firms export, the removal of trade barriers increases the size of exporters and makes less productive firms shrink or even die. Consequently the reallocation process has a positive effect on aggregate productivity and, naturally, on welfare. After liberalization, workers can enjoy higher wages and lower prices because a larger share of production is concentrated in more productive firms.

The point I want to address is that the welfare improvements of the model are actually due to the particular shape of the aggregate labor demand curve. In the model, this curve slopes upward\(^1\) in the employment-wage space. This implies that any movement of the curve from the right to the left, which occurs under trade liberalization, leads to an increase in wages and, if labor supply is elastic, in employment as well.

An upward-sloping aggregate labor demand curve makes assumptions about the nature of the firm’s fixed production costs. In the Melitz model, this cost is measured in units of labor, implying that it increases when the wage increases. Once free entry and exit are taken into account, equilibrium profits are on average proportional to fixed costs. As a consequence, in equilibrium, if the wage increases, profits have to increase too to satisfy the entry and exit conditions. Given returns to scale are increasing, in order to obtain higher profits, firms have to increase their size. This leads to an upward sloping aggregate labor demand curve\(^2\).

I illustrate below the case of an economy that differs from Melitz's in terms of fixed costs. In this economy, the costs do not represent labor, but aggregate output, that is, they do not increase when the wage increases. I show that by changing this dimension the aggregate labor demand curve may actually be decreasing. This implies that any movement of the curve from the right to the left no longer leads to an increase in wages. The wage decreases and given elastic labor supply employment decreases. More importantly, this generates a reduction in aggregate output and welfare.

The choice between these two assumptions is often considered as a matter of convenience in the literature\(^3\). I show it implies two different stories

---

1Here the term 'aggregate' is important. Labor demand at the firm level is a decreasing function of the wage, but once we take into account the process of firm entry, the resulting aggregate labor demand curve is upward sloping.

2See Benhabib and Farmer (1994) for a discussion on the slope of the aggregate labor demand curve in models with monopolistic competition and aggregate demand externalities.

3There are many examples of models that have considered that fixed costs are measured
about the effect of trade on welfare in a context of firm selection without scale effects. The first states that increases in the productivity threshold following liberalization improves welfare because output is biased toward the most productive firms (Melitz 2003). The second story suggests that this bias can actually be harmful in terms of welfare. The welfare implications of trade liberalization have been questioned in the empirical literature. In particular, some papers have assessed the presence/absence of scale effect (Head and Ries 1999, Trefler 2004). My comment can be seen as providing a formalization for this debate.

2 One model, two different assumptions

I consider the closed-economy version of Melitz (2003). The exposition of the model is nevertheless different in that, instead of considering the wage as the reference price, I normalize another price so as to highlight the mechanism taking place in the labor market. The reason why I only focus on the closed-economy framework is because trade liberalization in the Melitz model takes the form of an increase in aggregate productivity. Consequently, any shock increasing productivity in the closed economy has similar effects4.

Time is continuous. Welfare is derived from the consumption of an aggregate good, which is produced in quantities \( Q \) from a set of inputs \( Z \) according to the production function

\[
Q = \left[ M^\eta \int_{z \in Z} q(z)^{\frac{\eta+1}{\sigma}} dz \right]^{\frac{\sigma}{\sigma-1}},
\]

where \( z \) is a given variety of inputs consumed in quantities \( q(z) \) and \( M \) is the mass of varieties. \( \sigma > 1 \) is the elasticity of substitution between inputs and \( \eta > -1 \) influences the love for variety in the model, which is increasing in \( \eta \). For example, when \( \eta = 0 \), equation (1) reduces to the standard utility function with love for variety as in Melitz (2003) or Krugman (1980). When \( \eta = -\frac{1}{\sigma} \), the function is rescaled as in Blanchard and Giavazzi (2003) so as to neutralize any love for variety. In the extreme case where \( \eta = -1 \), only average consumption of varieties matters for the value of \( Q \) and a larger mass of varieties does not have any impact on \( Q \). The price of the aggregate good is normalized to one.

Under specification (1), the demand of input \( z \) writes as \( q(z) = Q M^{\eta \sigma} p(z)^{-\sigma} \).

Inputs are produced by firms which use labor as a factor of production and compete under a monopolistic framework. Labor is supplied inelastically in quantity \( L \). Firms differ in productivity. When productivity

in units of output. This is true for Hopenhayn (1992), the original paper on which the Melitz (2003) model is based. It has also been used extensively in models of the “new economic geography” (e.g., see Chapter 8 in Baldwin et al., 2003). Some macroeconomic models with heterogeneous firms such as Hopenhayn and Rogerson (1993) and Restuccia and Rogerson (2008) also consider this assumption. They differ from Melitz (2003) in that they assume perfect competition on the goods market which implies a downward sloping aggregate labor demand curve. Finally, some models of trade and unemployment also belong to this category, for example Egger and Kreickemeier (2007) and Janiak (2007). Felbermayr and Prat (2007) use this assumption, but rely on preferences that neutralize any external scale economies and their result on welfare is in line with Melitz (2003). See Section 4.1 for a discussion on this point.

4 An example of shock increasing aggregate productivity in the model is a decrease in the sunk entry cost denoted by \( c_e \) below. Results hold in the open-economy case, that is for a decrease in the fixed cost of exporting or the iceberg cost, and for an increase in the number of trading partners.
is equal to $\phi$, then a variety is produced in quantities $q(\phi) = \phi n(\phi)$ and profits are equal to $\pi(\phi) = r(\phi) - w n(\phi) - C$, where $r(\phi) = p(\phi) q(\phi)$ is revenue, $p(\phi)$ the fixed price, $w$ the wage, $n(\phi)$ firm-level (production) employment and $C$ is a fixed cost paid by the firm.

I compare two economies, which are called $A$ and $B$ and only differ in the nature of the fixed cost $C$ paid by a firm when producing. The economy $A$ is similar to Melitz’s: in this economy, the fixed cost takes the form of a minimum mass of labor which is required for production to take place. Hence, in the economy $A$, $C = c w$, where $c > 0$ is the labor requirement. In the economy $B$, the fixed cost represents a certain amount of the aggregate good. In this economy, $C = c$.

In order to enter the industry of inputs, a firm has to pay a sunk cost $C_e$. Like for the fixed cost $C$, this sunk cost takes the form of labor in the economy $A$ and aggregate output in $B$. In the former $C_e = c e w$ and in the latter $C_e = c e$. Once the sunk cost is paid, productivity is revealed. It is a draw from a cumulative distribution function $F$, with $f$ the associated density. I denote by $\phi^*$ the productivity threshold such that if a firm draws a productivity parameter higher than $\phi^*$ profits are positive and the firm chooses to stay in the industry. If the productivity draw is lower, then the firm decides not to enter. I denote by $M_e$ the mass of firms paying the sunk cost, which writes as $M_e = \frac{1}{1 - F(\phi^*)} M$ in steady state.

Finally, while producing, a given firm can be hit by a negative productivity shock with probability $\delta$ and is forced out of the industry.

3 Equilibrium

Under the above framework, the first-order conditions of a firm with productivity $\phi$ imply the following firm-level labor demand (excluding the fixed cost) and markup:

$$\frac{\sigma - 1}{\sigma} \phi^{\sigma-1} M^\frac{1}{\sigma} n(\phi)^\frac{1}{\sigma} = w$$

and

$$p(\phi) = \frac{\sigma}{\sigma - 1} w,$$

which allows to link the relevant firm-level variables for two firms with productivity $\phi_1$ and $\phi_2$:

$$\frac{p(\phi_1)}{p(\phi_2)} = \frac{\phi_2}{\phi_1}; \quad \frac{n(\phi_1)}{n(\phi_2)} = \left( \frac{\phi_1}{\phi_2} \right)^{\frac{1}{\sigma}}; \quad \frac{r(\phi_1)}{r(\phi_2)} = \left( \frac{\phi_1}{\phi_2} \right)^{\frac{\sigma-1}{\sigma}}. \tag{2}$$

In steady state, expected profits before entry have to be equal to the sunk entry cost and profits for a firm with productivity $\phi^*$ are zero. As in the Melitz model, these conditions respectively lead to the two relations

$$\pi_e = \frac{\delta C_e}{1 - F(\phi^*)} \tag{3}$$

and

$$\pi_e = C \left\{ \frac{\phi^* (\phi^*)^{\frac{\sigma-1}{\sigma}}}{\phi^*} - 1 \right\}, \tag{4}$$

where $\phi^* = \left( \int_{\phi^*}^\infty \phi^{\sigma-1} f(\phi) \frac{d\phi}{1 - F(\phi^*)} \right)^{\frac{1}{\sigma-1}}$ is average productivity, which is increasing in $\phi^*$.

In the case of the economy $B$, (3) and (4) give the equilibrium value of expected profits $\pi_e = \pi(\phi^*)$ and the productivity threshold $\phi^*$. In the case of the economy $A$, the costs $C$ and $C_e$ are endogenous, then (3) and
(4) give the ratio of profits to wage \( \frac{\pi_e}{w} \) and the threshold \( \phi^* \). Importantly, as I show in the next Section, these differences may produce different shapes of the aggregate labor demand curves. More importantly, welfare implications may also differ in the two economies.

On the other hand, notice that in both economies, conditions (3) and (4) jointly determine \( \phi^* \) and, therefore, also \( \phi_e \), independently from the wage. This is important for a clear understanding of the labor market equilibrium, which I describe below. In both economies, higher \( c \) or lower \( c_e \) is associated with higher \( \phi^* \).

Aggregate labor demand is

\[
N = Mn(\phi^*) + Mc + M_e c_e
\]

in the economy \( A \) and

\[
N = Mn(\phi^*)
\]

in the economy \( B \). Equations (5) and (6) tell us that in the economy \( A \) labor is allocated to several tasks, which are production, minimum employment requirement and investment in new varieties, while in the economy \( B \) employment is only allocated to production.

As standard in the literature, one can analyze the effect on wages to understand the effect on welfare. If wages increase, welfare increase too and vice-versa if they decrease.

4 Welfare implications

4.1 Love for variety

In this Section, I consider the case where \( \eta = 0 \), i.e. the economy displays a love for variety as in Melitz (2003). In this case, the equilibrium mass of firms is

\[
M = p(\phi^*)^{\sigma-1}
\]

in the two economies. Equation (7) states that the lower the price fixed by firms is on average, the smaller the mass of varieties is. The absence of competition effect is due the particular form of the function (1), which implies that markups are independent of the mass of firms in the economy. Together with (2)-(6), (7) allows to derive the aggregate labor demand functions in terms of \( \phi^* \) and \( w \), which is

\[
N = \left( \frac{\sigma}{\sigma - 1} w \right)^{\sigma-1} \left[ \frac{c}{\phi^*} \frac{\sigma - 1}{\sigma - 1} + \frac{c}{\phi^*} \frac{\sigma - 1}{\sigma - 1} + \frac{\delta}{1 - F(\phi^*)} \frac{\phi_e}{\phi^*} \right] (8)
\]

in the economy \( A \) and

\[
N = \left( \frac{\sigma}{\sigma - 1} w \right)^{\sigma-2} \frac{c}{\phi^*} \frac{\sigma - 1}{\sigma}
\]

in the economy \( B \). If the elasticity of substitution is lower than 2, then the aggregate labor demand curve is increasing in \( w \) in the economy \( A \) and decreasing in the economy \( B \). This implies that any movement of the curve from the right to the left leads to a different impact on wages. For instance, a shock increasing the productivity threshold \( \phi^* \) (e.g. due

\[
^5\text{In the next Section, I show one can increase the threshold on the elasticity of substitution by playing with the value of }\eta.\]
Figure 1: Labor market impact of an increase in $\phi^*$ in the two economies

Notes: in the two graphs, $S$ is the labor supply curve, $D$ and $D'$ are the initial and final labor demand curves respectively, $w$ is the wage and $N$ is employment.

6If labor supply is elastic, an increase in $\phi^*$ raises equilibrium employment in the economy $A$ and diminishes it in the economy $B$ (see Janiak 2007). Note also that in the case of the economy $A$, if labor supply is very elastic, the effect on employment may be negative too.

6
in the economy $A$ and

$$
N = \left( \frac{\sigma}{\sigma - 1} w \right) \frac{(\sigma - 2)\phi^{*}\sigma^{1-\eta}}{\phi^{*}\sigma^{1-\eta}(\sigma - 1)}
$$

in the economy $B$.

When $\eta > \frac{\sigma - 2}{\sigma - 1}$, the aggregate labor demand curve in the economy $B$ is decreasing. It is always increasing the economy $A$. In this range of parameter values, a shock increasing the productivity threshold $\phi^{*}$ leads to lower wages and welfare in the economy $B$ and welfare improvement in the economy $A$.

4.3 Decreasing external returns to scale

Suppose now the parameter $\eta$ takes value between $-1$ and $-\frac{1}{\sigma}$. The technology associated with the aggregate good then displays decreasing returns to scale in the number of varieties. In this case, the aggregate labor demand curves still write as in (11) and (12). However, the difference is that, for these values of $\eta$, the curve is now downward sloping in the economy $A$ and upward sloping in the economy $B$, leading to opposite effects of productivity-enhancing shocks.

4.4 Generalization

Figure 2: Labor market equilibria in the generalized economy

I consider now the generalized case of an economy where a share $\alpha$ of the fixed costs represents labor and a share $(1 - \alpha)$ is a certain amount of the aggregate good. In this economy, $C = \alpha c + (1 - \alpha)cw$ and $C_e = \alpha c_e + (1 - \alpha)c_e w$. If $\alpha = 1$, then the economy is identical to the economy $A$ previously described and if $\alpha = 0$ it is the same as the economy $B$. It
can be shown that the aggregate labor demand curve is then
\[ N = \alpha N^A(w) + (1 - \alpha)N^B(w), \]
where \( N^A(w) \) and \( N^B(w) \) are the labor demands described in equations (11) and (12).

Thus, aggregate demand is simply a linear combination of the aggregate demands in the two economies \( A \) and \( B \). If the curve had a positive slope in \( A \) and a negative one in \( B \), then the generalized economy may be characterized by multiple equilibria as depicted in Figure 2.

5 The aggregate labor demand curve in the empirical literature

The literature on empirical labor economics usually does not really aim to study the aggregate shape of the labor demand curve as a primary topic. Most of the studies adopt a micro perspective and are rather interested in the individual firm behavior. However, before working with firm-level data, those studies were relying on aggregate or industry-level datasets. Hamermesh (1996) provides an excellent review of this literature. See for instance the papers by Berndt and Khaled (1979), Berndt and Wood (1975), Chung (1987), Diewert and Wales (1987), Griffin and Gregory (1976), Magnus (1979), McElroy (1987), Morrison (1986, 1988), Pindyck (1979), Pindyck and Rotemberg (1983) and Segerson and Mount (1985), among others. All those studies rely on aggregate data or data at the large-industry level. They also take into account simultaneity issues between supply and demand in the estimation procedure. They may so provide a description of the aggregate labor demand curve. All of them actually report negative wage elasticity of the demand for labor, suggesting that this curve slopes downward.

Moreover, other papers have tried to study the behavior of firm entry and exit following wage shocks. These analysis have unfortunately not been included into the whole estimation of an aggregate labor demand curve. Anyway they produce results which are in line with a downward-sloping labor demand curve. Some estimate the probability of plant closing in terms of the wage level; see for instance Hamermesh (1988, 1996) and the references therein. They predict that increase an in wages leads to greater plant closing. Others have analyzed the role of wages in plant openings and show that high wages tend to deter entry; see Hamermesh (1996) for a review.

6 Conclusion

Is aggregate labor demand an upward- or a downward-sloping curve? Most labor economists would argue that labor demand decreases when the wage is higher. But, when an economy is characterized by increasing returns to scale, the opposite may be the case. Only empirical studies can answer this question.

As this comment has illustrated, the slope of the aggregate labor demand curve has important implications in the Melitz model. When demand is an increasing function of the wage, an increase in the productivity threshold has a positive effect on welfare, but has a negative impact when demand is a downward-sloping curve. The shape of the curve depends on
several assumptions such as the nature of the fixed cost, the elasticity of substitution and the love for variety.

For instance, with the standard CES utility function, when the fixed cost represents a minimum mass of labor which is required for production start up, as in Melitz (2003), aggregate labor demand is an upward-sloping curve. In this case, productivity-enhancing shocks are welfare improving. On the other hand, if one considers a different assumption about the nature of fixed costs, the results may be inverted. This is the case when the fixed cost is measured in units of aggregate output instead of labor.

Several conclusions can be drawn from this comment, depending on the particular assumptions one makes. Firstly, if it appears that reallocations are welfare improving and the aggregate labor demand curve slopes downward, a first interpretation of my results is that the Melitz model fails to reproduce an aggregate labor demand curve consistent with the empirical evidence. In this case, the Melitz and Ottaviano (2008) model may be more relevant as it introduces competitive effects that may invert the shape of the labor demand curve and provide welfare-improving properties from reallocation shocks. Secondly, if we believe in the Melitz model but think the aggregate labor demand curve slopes downward (as illustrated by empirical studies), a second conclusion is that policy makers should pay attention to the competitive structure specific to each industry before liberalization 

Finally, empirical studies of aggregate labor demand need to take account of firm entry and exit in their analysis, which could bias the estimates. Further research should help in resolving these puzzles.

References


\footnote{See Janiak (2007) for an application of this argument to employment issues.}


