Notes on BEJK 2003, by Mike Waugh

BEJK (Bernard, Eaton, Jensen, and Kortum (2003)) is an important paper in the cannon of the EK environment. Essentially, they extend the Ricardian environment in one important direction, i.e., they allow for imperfect competition with variable markups rather then competitive marginal cost pricing of the Eaton and Kortum (2002) setting. Now this extension was **not** just a modeling exercise, but it was done to address some important facts / measurement issues that Bernard and Jensen (and others) had been documenting prior to this paper.

A second contribution, is that in this paper they are very explicit about how to think of their framework as DGP (data generating process) at the micro-level. In other words, the functional forms they introduced are not just to facilitate aggregation with paper and pencil, but provide a simulation based method to examine and study outcomes at the micro-level. This and the EKK paper were key innovators on this and helped set the stage for how Ina and I in Simonovska and Waugh (2014) thought about these models.

1. A Quick Primer on the Measurement of Productivity.

How do we measure productivity? Turns out even the best economists continually trip over this question. So let's review some concepts. First, let's think of a simple production function like this

$$y_i(\omega) = z_i(\omega)\, n_i(\omega). \tag{1}$$

where we have the z term and then the labor term n. Often we say things like z is productivity. For example, in the Melitz (2003) world we might start to say things like, well exporters are high z firms, therefore exporters are more productive. Or in the Eaton and Kortum (2002) world, we say, well we are buying things from (in a likelihood sense) more productive producers. But we need to be careful. These are fine statements about the z, but this is **not** what is measured in the data when someone talks about productivity.

The issue is that those statements above are sometimes referred to as **physical** productivity or Hsieh and Klenow (2009) developed some terminology referring to this as **TFPQ** (total factor productivity measured in quantity units). But we (alomst) never see physical productivity because we are always measuring things with prices attached to them.[1] What we more often see in the data is something like this which I will call value added per worker:

$$\frac{p_i(\omega) z_i(\omega)\, n_i(\omega)}{n_i(\omega)} = p_i(\omega) z_i(\omega) \tag{2}$$

[1]This observation motivates why IO economist focus on things like concrete or cardboard boxes where the goods are homogenous and we might have some ability to get at physical productivity.

or how many dollars of output is produced per unit of labor.Hsieh and Klenow (2009) call **TFPR** (total factor productivity measured in revenue units). At this point, you may say "no big deal." Well let's impose some kind of pricing protocol like in Eaton and Kortum (2002), then notice that value added per worker is

$$p_i(\omega)z_i(\omega) = \frac{w_i}{z_i(\omega)} \times z_i(\omega) = w_i. \tag{3}$$

So even though producers are all heterogenous in their physical productivity, value added per worker is the same and equal to the wage. It's worth thinking about the intuition here, what is happening is that the theory of value (i.e. how ps are determined) mean that high z goods are *valued* with low ps and in the exact same proportion. Then when we think about how much value added is coming from high z producers, well by the same amount as low z producers.

These kind of observations are really important because how they help us interpret / think through the data. As we will discuss below, there is a lot of dispersion in value added per worker across firms / plants. This then begs for some kind of deviation from what we discussed above. These kinds of observations had huge impacts in fields outside of trade. The Hsieh and Klenow (2009) misallocation paper plays all off of these insights and was a motivation for "wedges" to be thing. My own work in Gollin, Lagakos, and Waugh (2013) makes similar observations when thinking about differences in value added per worker across sectors in developing countries, i.e. value added per worker in agriculture is much lower than value added per worker outside of agriculture and this is suggestive of misallocation.

The most obvious deviation is to say that prices do not just reflect marginal costs, but also markups. But its important to note that this is not sufficient. Consider the CES + monopolistic competition world of Melitz (2003). And lets measure value added per worker again

$$p_i(\omega)z_i(\omega) = \frac{\mu\, w_i}{z_i(\omega)} \times z_i(\omega) = \mu\, w_i, \tag{4}$$

where μ is the markup. This did not solve any issue at all. Value added per worker is still equated across firms, and the constant markup is behind this. This is a key issue with the CES preference structure and monopolistic competition structure. Also viewed through this lens, a lot of very loose claims about the importance of firms in the Melitz (2003) framework need to be carefully scrutinized because the largest, exporting firms value added per worker is the same as the smallest (near exit threshold) firms.[2]

[2] One out for Melitz (2003) is how fixed costs are treated in the data. Even if some of the fixed costs show up in value added per worker, then value added per worker will start to vary across firms.

2. BEJK Facts.

A key motivation of the BEJK paper are a couple of facts about firms / plants in trade.[3] I'm going to discuss them here. Note the data they have is old — this would be something to turn an AI tool on and update these facts.

1. Exporting behavior is very selected, most firms don't export. The BEJK data is old, but they get something like only 21 percent export anything. Think about this like the extensive margin. Then if they do export, they don't export much relative to their output — 2/3rds only sell less than 10 percent of the output abroad.

2. Yet, nearly 15 percent of U.S. manufacturing output is exported. That's strange I just told you that at the firm level, not much is exported. Well the issue is that those that are exporting are huge. Exporting plants are about 5 times bigger than non-exporters.

3. Facts [1.] and [2.] have little to do with industry. I think they are arguing against older view of trade as being driven by properties of an industry (e.g. capital intensity etc.).

- Note that behind these facts [1.] and [2.] lies the motivation in Melitz (2003) — we need something that will generate selection into exporting like the data. Melitz (2003) did it with heterogeneous physical productivity and fixed cost of exporting. However, Melitz (2003) faces the following problem in the next fact.

4. **Measured** productivity differs substantially across firms. And **measured** productivity is larger for exporters relative to not exporters. See Figure 2 or Table 2. For example, even within narrow industries, exporters have **measured** productivity of 10 percent larger relative to non-exporters. Note how I keep emphasizing **measured**. As discussed above, the competitive Eaton and Kortum (2002) can not reconcile this (there is a separate issue that there is no concept of a firm either). CES and Melitz (2003) can not either.

This is the launching point of the paper. They want a model that generates both selection into exporting **and** measured productivity dispersion. Their solution is to introduce Bertrand competition with firms competing head-to-head on price, markups become variable and depend on the gap between a firm's productivity and its nearest competitor's. This breaks the measured productivity equalization result and delivers the facts above.

3. The BEJK Environment

It's essentially the same as in Eaton and Kortum (2002), so I keep it brief and highlight certain departures that BEJK pursue.

[3]Firms and plants are not necessarily the same thing because of multi-establishment firms. It's true most firms are one plant firms, but also the largest firms are likely to multi-establishment.

Commodity space: Goods are indexed by $\omega \in [0,1]$. There are J countries. Many firms in each country can produce good ω, so each country's "version" of ω is not differentiated — they are perfect substitutes.

Technologies and trade frictions. In each in country i, there are many **firms** producing a good ω and each firm has a labor only technology defined by the productivity level $z_{ki}(\omega)$. Notice the k here. The subscript k indicates that this is the kth best firm producing good ω in country i. This is distinct relative to Eaton and Kortum (2002) as in that environment all k producers in country i have access to the best technology. Here some firm has a best technology, $k=1$, second best, $k=2$, and so forth.

To summarize the kth producer has the production function:

$$y_{k,i}(\omega) = z_{ki}(\omega) n_{ki}(\omega), \tag{5}$$

where labor, $n_i(\omega)$, is the only factor of production.

As usual, frictions to trade are modeled as iceberg trade costs. Again, I use the EK notation where the first subscript is the importer, the second sub-script is the exporter.

Consumers. In each country, there is a representative consumer with preferences:

$$U_n = \int_0^1 \left\{ \sum_{j,k} x_{k,nj}(\omega)\, u(c_{k,nj}(\omega)) \right\} d\omega. \tag{6}$$

where $x_{k,nj}(\omega)$ is an indicator function that takes the value one if good ω is sourced from the kth producer in j and zero for all other destinations. Again, I'm emphasizing the view that within ω, all j, k, goods are viewed as perfect substitutes.

It's worth pausing here to reflect on the perfect substitutes assumption. It is different from a similar model in Atkeson and Burstein (2008). That model is one of the "fruit salad" so the consumer in country n wants to consume in positive amounts all j and k. This is distinction consequential in how markups are determined. Atkeson and Burstein (2007) have a short note on this if you want to see more. I'll say more about this later.

Endowments. Finally, each country has a labor endowment of mass of N_i and the representative consumer in each country supplies his labor inelastically.

Distribution of Technologies. This is kind of the wild part. Now rather than specifying the distribution for the best technology, we need to know more. Specifically, we need to specify a joint distribution between the z_{ki} within country i across the k firms. In other words, what is the joint distribution between the best firm z_{1i}, z_{2i}, ..., etc. Turns out that BEJK has a nice distribution for that too, which we will talk about later.

3.1. Firm pricing

This is everything. Let's walk through this carefully.

Step 1. Because of the perfect substitutes assumption. One firm takes the whole market in n for good ω. This just follows from our choice rule (which we could derive as before in the EK setting):

$$x_{k,nj}(\omega) = \begin{cases} 1, \quad \text{if} \quad p_{k,nj}(\omega) \quad \leq \quad \min\limits_{k',j'} \left\{ p_{k',nj'}(\omega) \right\} \\ \\ 0, \quad \text{otherwise.} \end{cases} \tag{7}$$

Step 2. Who actually wins? Well all firms could set prices equal to marginal costs and earn zero profits. Let's start with that, so

$$x_{k,nj}(\omega) = \begin{cases} 1, \quad \text{if} \quad \frac{w_j d_{nj}}{z_{kj}(\omega)} \quad \leq \quad \min\limits_{k',j'} \left\{ \frac{w_{j'} d_{nj'}}{z_{k'j'}(\omega)} \right\} \\ \\ 0, \quad \text{otherwise.} \end{cases} \tag{8}$$

We also know because of the ordering on k, that $k = 1$ will always win (thus all $x_{k,nj}(\omega) = 0, \forall k \neq 1$). And which country is the winner depends on the productivity advantage relative to the wage rate and trade costs in that location. This identifies the winner. Again, this is the same as in EK.

Step 3. At what prices? Here is the key thing, the winner knows he has market power and can set the price just to the point such that he keeps all other competitors out. That is so the inequality in binds. However, the winner also knows that he faces a downward sloping demand curve and does not want to choke off demand too much. With a CES demand curve, the winner will not want to go beyond the monopoly markup of $\frac{\sigma}{\sigma-1}$. Call the winner $j(\omega)^*$. Then given these arguments we have that

$$p_{1,nj^*}(\omega) = \min \left\{ \frac{\sigma}{\sigma - 1} \cdot \frac{w_{j^*} d_{nj^*}}{z_{1j^*}(\omega)} \, , \, \min_{k',j' \neq j^*} \left\{ \frac{w_{j'} d_{nj'}}{z_{k'j'}(\omega)} \right\} \right\} \tag{9}$$

so the next lowest price. Now there is a subtlety here BEJK want to emphasize. They break out this min operator by making the observation that the next lowest price can only be one of two occurrences: (i) the second best domestic competitor in j^* or (ii) the $k = 1$ best competitor elsewhere. So they rewrite this as

$$p_{1,nj^*}(\omega) = \min\left\{\frac{\sigma}{\sigma-1}\cdot\frac{w_{j^*}d_{nj^*}}{z_{1j^*}(\omega)}\ ,\ \min\left\{\frac{w_{j^*}d_{nj^*}}{z_{2j^*}(\omega)}\ ,\ \min_{j'\neq j^*}\left\{\frac{w_{j'}d_{nj'}}{z_{1j'}(\omega)}\right\}\right\}\right\}. \tag{10}$$

So to summarize: set the price either at the monopoly price or to keep the next best guy out. And the next best guy are either my own local competitor or international competitors from other sources.

Let me briefly return to the discussion of Atkeson and Burstein (2008). Again, their model is one of fruit salad and, in some ways, the pricing argument is simpler. All k, j firms are active because we want the variety. Then when these different firms set their price, they simply differentiate their demand curve taking into account the impact that they have on other firms in market n, ω. And the markup depends upon something like the shape of the demand curve and then something about how large each firm is in the market. In this setting, the markup is explicitly determined so that the best firm captures the entire market.

3.2. Markups and Measured Productivity

Now let's be precise about markups and connect back to our earlier discussion of measured productivity. In this discussion below, I'm going to **assume that we are in a symmetric world.** So that wages $w_1 = w_2 = w_J = 1$ and all trade costs are symmetric to illustrate things more clearly.

Define the markup for the winning firm as the ratio of price to marginal cost:

$$\mu_{nj^*}(\omega) = \frac{p_{1,nj^*}(\omega)}{\frac{d_{nj^*}}{z_{1j^*}(\omega)}}. \tag{11}$$

From our pricing rule, the markup is

$$\mu_{nj^*}(\omega) = \min\left\{\frac{\sigma}{\sigma-1},\quad \frac{z_{1j^*}(\omega)}{\tilde{z}_{nj^*}(\omega)}\right\} \tag{12}$$

where $\tilde{z}_{nj^*}(\omega)$ is the "effective productivity" of the next-best competitor—either the second-best domestic firm or the best foreign firm, adjusted for trade costs. Specifically,

$$\tilde{z}_{nj^*}(\omega) = \max\left\{z_{2j^*}(\omega),\quad \max_{j'\neq j^*}\{z_{1j'}(\omega)\}\right\}. \tag{13}$$

The key observation is that the markup depends on the ratio $z_{1j^*}/\tilde{z}_{nj^*}$ or the "gap" between the winner and the next-best competitor. When this gap is large, the winner has a substantial cost advantage and can charge a high markup without losing the market. When the gap is small, competition is fierce and the markup is driven toward one. All in all, markups will be varying across ω because different suppliers have different next-best competitors.

This observation is exactly what generates measured productivity dispersion. Recall that measured productivity (TFPR) is revenue per unit of input. For the winning firm selling in market n:

$$\text{TFPR}_{nj^*}(\omega) = p_{1,nj^*}(\omega) \times z_{1j^*}(\omega) = \mu_{nj^*}(\omega) \times d_{nj^*}. \tag{14}$$

where I've used the normalization that wages are equal to one. Now we see that $\text{TFPR}_{nj^*}(\omega)$ will be varying across ω because markups are varying with ω.

The not obvious question is about how $\text{TFPR}_{nj^*}(\omega)$ would vary with exporting behavior. We know that if a firm is likely to export, it's likely to have a large z. This is just natural Ricardian selection that arises in the EK world. But does it have high TFPR? That right now is not obvious to me. The key issue essentially is the following: conditional one being very good, what is the likely-hood that I also have a large gap versus my second best competitor.

4. Welfare Effects with Markups

Before moving onto the distributional assumption and then quantitative results. I want to walk through how markups might matter for the gains from trade. This will also help us understand some surprising results of BEJK.

Firms set prices as a markup over marginal cost:

$$p_{nk}(\omega) = \mu_{nk}(\omega; d, w) \cdot \frac{w_k d_{nk}}{z_k(\omega)}, \tag{15}$$

where $\mu_{nk}(\omega)$ is a markup that may depend on trade costs, wages, or other equilibrium objects. I'm not going to commit to a specific reason / market structure behind these markups, just that they might depend upon primitives and equilibrium objects. In this sense, what I'm doing is more general than the BEJK environment discussed above.

Budget constraint. With markups come profits — and we need to account for them somehow. For now, we will just have that profits accruing to households and, thus the representative consumer in country n faces the following budget constraint.

$$w_n N_n + \Pi_n \geq \int_0^1 \sum_j x_{nj}(\omega) p_{nj}(\omega) c_{nj}(\omega)\, d\omega, \tag{16}$$

where Π_n denotes total profits from country n's firms.

4.1. Welfare Effects of a Change in Trade Costs

Now I'm going to make very similar arguments as in the EK-competitive setting. We start with the Lagrangian and then totally differentiate it. So the consumer's problem gives rise to the Lagrangian:

$$\mathcal{L} = \int_0^1 \sum_j x_{nj}(\omega) u(c_{nj}(\omega))\, d\omega + \lambda_n \left\{ w_n N_n + \Pi_n - \int_0^1 \sum_j x_{nj}(\omega) p_{nj}(\omega) c_{nj}(\omega)\, d\omega \right\}. \tag{17}$$

And now we take the total derivative, he welfare effect of a change in d_{nj} decomposes as:

$$\frac{d\mathcal{L}}{d\, d_{nj}} = \underbrace{\int_0^1 \sum_k \frac{\partial \mathcal{L}}{\partial p_{nk}(\omega)} \frac{dp_{nk}(\omega)}{d\, d_{nj}} d\omega}_{\text{through prices}} \quad + \quad \underbrace{\int_0^1 \sum_k \frac{\partial \mathcal{L}}{\partial c_{nk}(\omega)} \frac{dc_{nk}(\omega)}{d\, d_{nj}} d\omega}_{\text{through quantities}} \tag{18}$$

$$+ \underbrace{\int_0^1 \sum_k \frac{\partial \mathcal{L}}{\partial x_{nk}(\omega)} \frac{dx_{nk}(\omega)}{d\, d_{nj}} d\omega}_{\text{through sourcing}} \quad + \quad \underbrace{\lambda_n \frac{d\Pi_n}{d\, d_{nj}}}_{\text{through profits}} \quad . \tag{19}$$

And then the same exact envelope type argument applies to the "through quantities" term and the "through sourcing" term and, thus these terms zero out. What is left is

$$\frac{d\mathcal{L}}{d\, d_{nj}} = \underbrace{\int_0^1 \sum_k \frac{\partial \mathcal{L}}{\partial p_{nk}(\omega)} \frac{dp_{nk}(\omega)}{d\, d_{nj}} d\omega}_{\text{through prices}} + \underbrace{\lambda_n \frac{d\Pi_n}{d\, d_{nj}}}_{\text{through profits}} \tag{20}$$

or direct effects through prices and profits.

Now lets open up the price effects term. First, recall that the partial derivative of the Lagrangian with respect to prices is

$$\frac{\partial \mathcal{L}}{\partial p_{nk}(\omega)} = -\lambda_n x_{nk}(\omega) c_{nk}(\omega). \tag{21}$$

The total derivative of the price satisfies:

$$\frac{d \ln p_{nk}(\omega)}{d \ln d_{nj}} = \frac{d \ln \mu_{nk}(\omega)}{d \ln d_{nj}} + \frac{d \ln w_k}{d \ln d_{nj}} + \mathbf{1}_{k=j}, \tag{22}$$

where $\mathbf{1}_{k=j}$ is an indicator equal to one if $k = j$ and zero otherwise.

Then the goal is to insert this into the first term of (20). To do so, define the expenditure-weighted average markup elasticity:

$$\bar{\varepsilon}^{\mu}_{nk,j} \equiv \frac{1}{X_{nk}} \int_0^1 x_{nk}(\omega) p_{nk}(\omega) c_{nk}(\omega) \frac{d \ln \mu_{nk}(\omega)}{d \ln d_{nj}} d\omega, \tag{23}$$

where X_{nk} is total expenditure by country n on goods from country k. Then after converting everything to the correct units we have

$$\frac{d \ln W_n}{d \ln d_{nj}} = -\pi_{nj} - \sum_k \pi_{nk} \frac{d \ln w_k}{d \ln d_{nj}} - \sum_k \pi_{nk} \bar{\varepsilon}^{\mu}_{nk,j} + \frac{\Pi_n}{E_n} \frac{d \ln \Pi_n}{d \ln d_{nj}}, \tag{24}$$

where $\pi_{nk} = X_{nk}/E_n$ is the expenditure share and $E_n = w_n N_n + \Pi_n$ is total expenditure.

What is going on here? Equation (24) has (i) the direct effect which is what would be the only thing present from the perspective of planner (ii) the GE-terms of trade effects. And then new terms (iii) how "average markups change" appropriately weighted and then (iv) how profits change. These last two terms are they key elements as to **why** trade economists would be interested in situations with variable markups. In particular, the third term is what a lot of trade economists would term "pro-competitive effects" of trade with the idea that lower trade costs **could** result in lower markups and, thus larger gains from trade.

A several points about this...

- First, notice that in a constant markup world both the elasticity term and the profit term would be zero. So CES + monopolistic competition is no-go for pro-competitive effects.

- The elasticity term is a complicated object. It is not obvious that the competitive effects are actually beneficial. They could go the other way. BEJK actually discuss related issues in depth when discussion how several different forces can move to offset things. As one example, very productive exporters may have their advantage reinforced by the reduction trade costs leading to a higher markup. So how do we know if the gap in (13) is increasing or decreasing?

  Moreover, what matters is not a simple average but a weighted average. So instances where most markups don't change, but for goods that have high expenditures they do, then this term might be more active. I need to dig into this more but this is very closely related to pass-through and it is a broader issue than trade, connects with monetary transmission etc.

- The profit term is interesting too, I think often overlooked. Notice that profits are connected to markups, so how the two new terms balance out is delicate. Profits also raise a question about who gets them and they start begging for heterogeneous consumers who have different exposures to labor vs. profit / capital income and how they change. Later we might discuss this as well.

References

ATKESON, A. AND A. BURSTEIN (2007): "Pricing-to-market in a Ricardian Model of International Trade," *American Economic Review*, 97, 362–367.

——— (2008): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 98, 1998–2031.

BERNARD, A. B., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): "Plants and productivity in international trade," *American economic review*, 93, 1268–1290.

EATON, J. AND S. KORTUM (2002): "Technology, geography, and trade," *Econometrica*, 70, 1741–1779.

GOLLIN, D., D. LAGAKOS, AND M. E. WAUGH (2013): "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*.

HSIEH, C.-T. AND P. J. KLENOW (2009): "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 124, 1403–1448.

MELITZ, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71, 1695–1725.

SIMONOVSKA, I. AND M. E. WAUGH (2014): "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 92, 34–50.