Putting Per-Capita Income back into Trade Theory

James R. Markusen
University College Dublin
University of Colorado, Boulder

Abstract

A major role for per-capita income in international trade, as opposed to simply country size, was persuasively advanced by Linder (1961). Yet this crucial element of Linder’s story was abandon by most later trade economists in favor of the analytically-tractable but counter-empirical assumption that all countries share identical and homothetic preferences. This paper collects and unifies a number of disjoint points in the existing literature and builds further on them using simple and tractable alternative preferences. Adding non-homothetic preferences to a traditional models helps explain such diverse phenomenon as growing wage gaps, the mystery of the missing trade, home bias in consumption, and the role of intra-country income distribution, solely from the demand side of general equilibrium. With imperfect competition, we can explain higher markups and higher price levels in higher per-capita income countries, and the puzzle that gravity equations show a positive dependence of trade on per-capita incomes, aggregate income held constant. In all cases, the effects of growth are quite different depending on whether it is growth in productivity or through neutral factor-endowment growth. The paper concludes with some suggestions for calibration, estimation, and gravity equations.

This Draft, May 5, 2010

Comments, suggestions and added citations welcome
1. Introduction

All international trade economists understand that many things can cause trade. However, our models and empirical analyses typically and appropriately tend to focus on one cause of trade at a time in order to understand how a particular basis for trade contributes to explaining trade patterns, determines gains from trade, and impacts on income distribution. That having been said, it seems that most trade theory focuses on production-side determinants of trade. It is typically assumed that consumers have identical and homothetic preferences within and across countries. Aggregate demand depends only on commodity prices and aggregate income, and it is independent of the distribution of income. I also believe that it is appropriate to suggest that no one thinks that this is a good empirical assumption and that it is made for analytical convenience and tractability.¹

If we control for differences in prices across countries, the observation of different budget shares can indicate either that preferences differ and/or that they are non-homothetic. Two pure cases can be distinguished: one in which countries have homothetic but non-identical preferences and one in which countries have identical but non-homothetic preferences. I feel much more comfortable with the second alternative. Then demand differences are not only systematic but the hypothesis is testable and falsifiable.

The purpose of this paper is collect, synthesize, and build on fragmented results from existing research in order to offer a generic model of identical but non-homothetic preferences in order to present a unified and testable set of results. In section two, the preferences are presented and analyzed and then placed on top of a standard two-good, two-factor, two-country Heckscher-Ohlin model. The resulting model offers alternative explanations for such diverse phenomenon as growing wage gaps, the mystery of the missing trade, home bias in consumption, and a role for intra-country income distribution solely from the demand side of general equilibrium. In section 3, I add scale economies and imperfect competition and show that the model can offer alternative explanations for higher price levels and higher markups in high-productivity economies, and a higher trade volume between identical high per-capita income countries, aggregate income held constant. In both competitive and imperfect-competition cases the effects of growth are quite different depending on whether it is growth in productivity or in factor accumulation.

Before proceeding, a short literature review is in order and I should first point out that many of the ideas here are found in earlier papers focusing on specific issues. Several papers, obviously beginning with Linder, focused on monopolistic competition and the impact of non-homothetic preferences on intra versus inter-industry trade. Papers by Markusen (1986), Bergstrand (1990), and Francois and Kaplan (1996) draw out implications for intra-industry and total trade volumes. Matsuyama (2000) uses a competitive Ricardian model in which the South’s comparative-advantage goods are low-income-elasticity-of-demand goods to derive results similar to some here. Fieler (2010) uses a Ricardian approach

¹A more egregious assumption is made in much of the strategic trade-policy literature: quasi-linear preferences in which the income elasticity of demand for increasing-returns good(s) is zero. Yet the industries offered as examples, such as aircraft and electronics are surely goods with income elasticities greater than one! The present author has of course been guilty himself of this atrocity.
ala Eaton and Kortum (2002), which yields outcomes related to those from monopolistic competition. There is a dispersion of technologies across goods and a variability of labor efficiency across countries. High income elasticity goods have a higher dispersion and are produced in high-income countries. This higher dispersion leads to more trade among the high-income countries relative to low-income countries.

Markusen (1986) is a three-region model with “east-west” trade among identical high-income (north) countries and “north-south” trade between the high and low per-capita-income countries. High income-elasticity goods are both capital-intensive and differentiated, while the homogeneous low -income elasticity good is labor intensive. The volume of east-west (north-north) trade is then higher than north-south trade relative to identical and homothetic preferences. Fieler (2010) makes substantial theoretical progress on this sort of multi-country prediction in her Ricardian model and deals with south-south trade as well.

Table 1 gives some income-elasticity estimates from broad categories of consumption goods from Markusen and Hunter (1989). Non-homotheticity similarly means that the shares of a country’s national expenditure for goods vary with levels of per-capita income. Table 1 shows shares from Hunter (1991) and Cassing and Nishioka (2009). Simply “eyeballing” shares in these two papers suggests that preferences are not just differing randomly across countries but observed shares are related to per-capita income.

Hunter (1991) shows that the influence of non-homotheticity is in the direction of reducing the volume of trade. A counter-factual analysis neutralizes the estimated non-homotheticity and finds that the effect of imposing homotheticity is to raise trade flows by 29 percent. Cassing and Nishioka (2009) use a neutralization exercise similar to Hunter’s and find that developing countries consume relatively more labor-intensive goods than under preference homogeneity. Second, they find that preference biases between rich and poor countries explain a larger proportion of missing factor trade than do differences in technology, though preference differences are not distinguished from non-homotheticity. Fieler’s (2010) empirical results are strong and convincing. Her model correctly predicts both the large volume of trade among high-income countries and the low volume of trade among poor countries, whereas previous gravity-type work has done a poor job on the latter.

These results contrast with Bowen, Leamer and Sveikauskas (1987) and Trefler (1995) who do introduce non-homothetic preferences into their analyses and get weak value added from doing so. Neither paper addresses non-homotheticity as a cause of missing trade or home bias (Trefler does find it helps solve the “endowment paradox”). Reimer and Hertel (2010) find that non-homothetic preferences play only a small role in missing trade over broad categories of goods. Rather, they find that, as income grows, it is directed toward the relatively capital-intensive version of a given good.

Results in the section introducing imperfect competition are similar to those found in Wong (2003), Hummels and Lugovskyy (2009) and Simonovska,(2009) which is that markups and hence the price level will be higher in the high per-capita-income country. Simonovska gets strong empirical support for this relationship, also found in earlier papers including Hsieh and Klendow (2007). I should also note that Simonovska carefully considers identical products, which eliminates quality issues which could be an alternative explanation for systematic price differences by per-capita income. Essentially the same result was found
by Wong for pricing of identical pharmaceutical products.

An area where per-capita income does play an important role is in the analysis of product quality. If a consumer is going to buy only one unit of a good or zero, then the quality demanded is likely to depend on per-capita income. This makes the average level of per-capita income important for trade but also the intra-country distribution of income matters for inter-country trade. Because the issues connected with product quality are somewhat better understood and because they require a quite different analytical approach, they will not feature in this paper despite their importance.2

2. A Generic Model

The preferences we will use are variation on a standard Stone-Geary utility function, to be introduced shortly. The production side of the model is deliberately Heckscher-Ohlin to permit an easy comparison with traditional results. There are two good (X and Y), two factors of production (K and L) and two countries home and foreign (h and f).

Throughout the paper, the following assumptions are made.

(1) good X is relatively capital intensive, and Y is relatively labor intensive
(2) good X has an income elasticity of demand greater than one
(3) good X is the increasing-returns good if there is one (section 3)
(4) the labor supply to production is identical to the number of households
(5) country h is relatively capital abundant when relative endowments differ
(6) country h has higher productivity when productivities differ across countries

Most of these assumptions are without loss of generality, but the intersection of (1) - (4) matters; in particular, that the capital-intensive good has the high income elasticity of demand. Empirical support for this assumption is found in Bergstrand (1990) and indirectly in Hunter (1991) and Cassing and Nishioka (2009). Matsuyama (2000) uses an equivalent assumption in his Ricardian model: the South’s comparative advantage goods are low-income-elasticity goods. Fieler’s (2010) theory is also Ricardian, and the theory and empirical evidence deliver a related result: the south has a comparative advantage in low-income elasticity standardized goods while the north has an advantage in high-income-elasticity goods whose production technologies are more variable across countries.

Since we will focus on a limited number of experiments, some short-hand terminology is used throughout. Productivity advantage or growth, or higher productivity refers to an equal proportional Hicks-neutral productivity advantage or growth in both sectors

in one country (always country h). Factor accumulation refers to a equal proportional
growth in both factor endowment of one or both countries: factor accumulation increases the
number of households in the same proportion.

Lower-case letters denote per-household quantities. In addition to x and y, there is a
parameter \( z > 0 \) at the household level. Preferences or utility \( u \) are given as follows.

\[
    u = (x + z)^\beta y^{1-\beta}
\]

We could interpret \( z \) as an endowment good and assume that households cannot buy or sell \( z \).
x could be televisions and \( z \) could be watching a sunset (non-rivaled and non-excludable:
sitting on a dock on the bay as Otis Redding might say). The assumption that \( x \) and \( z \) are
additive has little to with the results of this paper, but has the advantages that (a) there is a
simple analytical solution for demand and (b) aggregate demand does not depend on the
distribution of income (with a qualification noted below).\(^3\)

It is more common to see Stone-Geary written with \( (y - z) \) instead of \( (x + z) \), with \( z > 0 \) then referred to as a “minimum consumption requirement”. But this leads to a problem if
income is insufficient to purchase the minimum consumption requirement and no household
will ever be observed to purchase only good \( Y \). In addition, our formulation in (1) will mean
that the price elasticity of demand for \( X \) will be falling in per-capita income, a property
exploited in Wong and in Simonovska.

Let \( m^i \) denote the income of household \( i \) and let \( p_x \) and \( p_y \) denote the prices of \( X \) and \( Y \).
The households budget constraint is given by:

\[
    m^i = p_x x + p_y y
\]

Maximization of (1) subject to (2) gives the following Marshallian demand functions.

\[
    x^i = \max \left[ 0, \left( \beta - 1 \right) z + \frac{\beta m^i}{p_x} \right] \quad y^i = \min \left[ m^i, \frac{(1 - \beta)(m^i + p_x z)}{p_y} \right]
\]

\[
    x^i > 0 \quad \text{iff} \quad m^i > \frac{(1 - \beta)}{\beta} p_x z + m^0
\]

At low levels of income, the household buys only good \( Y \), and above the threshold income
indicated in (4) by \( m^0 \), begins to buy \( X \). This is an interesting and surely realistic point, and
it makes aggregate demand depend on the distribution of income. I will assume for much of

\(^3\)Virtually all the results go through with the CES-Cobb-Douglas version

\[
    u = (x^a + z^a)^{\frac{1}{a}} y^{1-\beta}
\]

except perfect aggregation: aggregate demand always depends on the
distribution of income. This function is hard to work with analytically (though easy for the
computer).
Properties of the preferences are illustrated in Figure 1, left panel. The Engels curve (prices constant) is given by \( OY/OA \): up to income \( m_0 \), given by (4) with equality, the household consumes only good \( Y \) and then has a constant marginal propensity to consumer \( X \) and \( Y \) as income rises. The share of \( X \) in consumption as household income rises is shown in the right-hand panel (calibration is \( \beta = 0.667 \)) of Figure 1.

Let \( L \) denote both the country’s labor supply and household measure so that \( Z = zL \) denotes the economy-wide “endowment” of \( Z \): \( z \) is a parameter, while \( Z \) is strictly proportional to the number of households. Given that (4) holds for all households, aggregate demand for \( X \) is independent of the distribution of income and given by

\[
X = \sum_{i=1}^{L} x^i = (\beta - 1)Z + \beta M \frac{p_x}{p_x} \quad Z = zL \quad M = \sum_{i=1}^{L} m^i \quad (5)
\]

Now consider the income elasticity of demand for \( X \) and assume that income grows through a productivity increase, holding the number of households \( L \) and therefore \( Z \) constant.

\[
\left. \frac{M}{X} \frac{dX}{dM} \right|_{Z \cdot M} = \frac{\beta M}{\beta M + (\beta - 1)p_x Z} = \frac{m}{m - m^0} > 1
\]

(growth through productivity improvement)

If growth instead occurs through neutral factor accumulation, \( Z \) is strictly proportional to \( M \), then

\[
\left. \frac{M}{X} \frac{dX}{dM} \right|_{Z \cdot M} = 1 \quad (growth \ through \ neutral \ factor \ accumulation) \quad (7)
\]

Using \( dX/dp_x = -\beta M/p_x^2 \), the Marshallian price elasticity, defined as positive, is

\[
\epsilon = -\left. \frac{p_x}{X} \frac{dX}{dp} \right| = \frac{\beta M}{\beta M + (\beta - 1)p_x Z} = \frac{m^a}{m^a - m^0} > 1
\]

(8)

Thus the per-capita income and the price elasticities of demand for \( X \) are (locally) the same, and illustrated in the right-hand panel of Figure 1.

---

\(^4\)The general n-good name for this type of demand function is “linear expenditure system” and is also used in Bowen, Leamer, and Sveikauskas, (1987) who refer to the \( Z \)s as “autonomous” expenditure. See Deaton and Muellbauer (1980) for an excellent general analysis.
An extension of this argument leads to the Prebisch (1950) - Singer (1950) hypothesis. If the countries differ in relative endowments (standard Heckscher-Ohlin), then neutral productivity growth in both countries will lead to a terms-of-trade deterioration for the labor-abundant country: the “south”.

The properties of aggregate demand for the economy, holding prices constant are shown in Figure 2. Let $Z_0$ denotes the initial value of $Z$. Hold $Z$ constant but allowing aggregate income to vary either through productivity or through capital accumulation, holding $L$ constant. This leads to an Engels income-consumption curve that starts at the origin and moves up the $Y$ axis (consuming only $Y$) to point $Y_0$ after which higher income will result in positive $X$ demand. At incomes above that which allows point $Y_0$ to be reached, the Engels curve is linear through A at income level $M_0$ and reaching B at income level $M_1$.

Consider point A and income level $M_0$ in Figure 2. Now suppose instead we let the economy grow through proportional factor accumulation, adding households in strict proportion to the increase in income, so $Z$ and $M$ grow to $Z_1$ and $M_1$, respectively. Now the Engels curve will be given by a ray through the origin and points A and C and aggregate demand is homothetic with respect to aggregate income. Figure 2 gives the first important result of the paper: a growing economy will look very different depending on whether growth is through productivity or capital accumulation on the one hand, or neutral factor accumulation on the other (aggregate income and households grow in strict proportion).

From this point on, I will some results in terms of simulations. All of the results are intuitive, some are found formally in earlier papers and I am quite sure that all of the qualitative properties of all results have no dependence on the specific parameters or other assumptions used in these specific examples. The initial “calibration” point is the one used in the right-hand panel of Figure 1: at productivity one, the income and price elasticity are 1.333 and the share of $X$ in consumption is 0.5; the value of $\beta = 2/3$ is used in this example and throughout the paper. As productivity grows without bound the income and price elasticity approaches one and the consumption share approaches its marginal value of 2/3 in Figure 1.

With the neutral and equal productivity growth in both sectors, the production frontier of the economy is growing radially, but demand is shifting toward good $X$. This generates a movement around the frontier toward $X$, so the relative price of $X$ rises as shown in Figure 3. But this generates the usual Stolper-Samuelson effect on relative factor prices, so the rental ($r$) - wage ($w$) ratio $r/w$ is rising as shown in Figure 3. Suppose we interpret capital as skilled labor or human capital and $L$ as unskilled labor. A neutral productivity growth generates an increase in the wage gap between skilled and unskilled labor. Thus we can get a wage gap phenomenon driven entirely by the demand side of the general-equilibrium model without appealing to trade or to skill-biased technical change.

Now consider differences in relative endowments, beginning with the two countries identical, under the assumption of costless trade. Move capital from $f$ to $h$ and labor from $h$ to $f$, implying the $Z$ rises in $f$ and falls in $h$ by an equal and opposite amount. Their Engels curves will move apart in Figure 2 but they remain parallel. Figure 4 shows the effect of widening the endowment differences. It graphs the share of world consumption and production in each country. The consumption shares in this exercise would be constant at 0.5

---

5An extension of this argument leads to the Prebisch (1950) - Singer (1950) hypothesis. If the countries differ in relative endowments (standard Heckscher-Ohlin), then neutral productivity growth in both countries will lead to a terms-of-trade deterioration for the labor-abundant country: the “south”.

---


under homothetic demand. But under our assumption that the capital intensive good is the high-income-elasticity good, the consumption shares are positively correlated with their respective good’s production share.

Figure 4 gives a demand-side explanation for two phenomenon that have previously been identified and attributed to production-side causes. The positive correlation between production and consumption shares has been one (of several) definition of “home bias”. Secondly, the volume of trade is less under our assumptions than is predicted under a standard Heckscher-Ohlin model and thus offers a demand-side explanation for the empirical puzzle of “missing trade” in Trefler’s (1995) terminology. The amount of missing trade is identified in Figure 4 and note that it continues to grow in importance once countries are specialized: production specialization cannot continue to increase but consumption specialization can. As noted earlier, non-homotheticity as a cause of missing trade was noted theoretically by Markusen (1986) is empirically verified in Hunter (1991) and is consistent with Cassing and Nishioka (2009). Closely related points in the Ricardian context are found in Matsuyama’s (2000) theory and in Fieler’s (2010) theoretical and empirical paper.6

As a final exercise with the competitive case, consider the role of intra-country income distribution which has been noted before.7 If each consumer in a country has enough income as given in (4) to want positive amounts of \( X \), then the linear property of the Engels curve means that redistribution of income within the country (subject to (4) continuing to hold for all households) does not affect aggregate demand. But if redistribution puts some households on the vertical section of the curve in Figure 1 or 2 where they only buy \( Y \) (points below \( Y^0 \)), then it does matter.

Let there be two sets of households, denoted with superscript \( p \) (poor) and \( r \) (rich). There are \( L' \) poor households and \( L' \) rich households, \( L = L' + L' \) with per-capital incomes \( m^p \) and \( m^r \) respectively. \( m^a \) will denote the average income. Assume that a household with average income would purchase positive amounts of \( X \) but poor households do not. With reference back to the minimum income condition in (4), we assume that

\[
m^a = \frac{m^p L' + m^r L'}{L}, \quad m^p < \frac{(1-\beta)}{\beta} p_x z = m^0 < m^a
\]

When there are just these two household types, only the rich ones will purchase \( X \). Suppose on the other hand, that all household types have the average per capita income. Aggregate demand \( X' \) and \( X'' \) in these two scenarios are given as follows.

---

6Markusen (1986) and Fieler (2010) explicitly explain large “north-north” versus small “north-south” (and south-south in Fieler) trade volumes. An entirely different explanation is advance in Markusen and Wigle (1990): north-south and south-south trade is small because the south is poor and because the pattern of wold protection discriminates heavily against north-south and south-south trade.

7 For example, Choi et al. (2009), Daglin et al. (2008), Francois and Kaplan (1996), Hallak (2006), and Mitra and Trindade (2005).
\[ X' = (\beta - 1)zL' + \frac{\beta m'L'}{p_x} \quad X^a = (\beta - 1)zL + \frac{\beta m^aL}{p_x} \]  

(10)

Subtract the second equation of (10) from the first, and substitute for \( m^a \) from the first equation of (9). The difference in aggregate demand is

\[ X' - X^a = (1 - \beta)zL^p - \frac{\beta m^pL^p}{p_x} > 0 \quad \frac{X' - X^a}{L} = \beta \left[ \frac{m^0 - m^p}{p_x} \right] \frac{L^p}{L} > 0 \]

(11)

where the inequalities holds by assumption (the income of poor households is too low to purchase \( X \)). Perfect aggregation does not hold with a wide distribution of household income and, for two countries with the same average income, aggregate demand for the luxury will be higher in the country with the more unequal distribution (those Mercedes in Africa).8

3. Imperfect competition, prices and markups

In this section, we add scale economies, imperfect competition, and free entry and exit of firms in the \( X \) industry in a standard model of Cournot competition, continuing with the assumption that \( X \) is a homogeneous good. \( Y \) is produced with constant returns under perfect competition. We assume segmented markets simply because the results are more interesting and in line with Simonovska for example, so the model is similar to Venables (1985) or Markusen and Venables (1988), the latter contrasting segmented and integrated markets cases.

To keep matter manageable, we will concentrate on the case of per-capita income differences arising from productivity difference, but the qualitative results are identical to the case where countries differ in endowment ratios. Suppose that there is just a single factor of production, \( L \), and distinguish between the household count and the “effective” or productivity-adjusted supply of labor in each of the two countries. One unit of \( Y \) uses one unit of effective labor supply and \( X \) uses \( c \) units of effective labor for marginal costs and \( F \) units for fixed costs. For the most part, we will assume that the total effective labor supply and hence total income measured in units of \( Y \) are the same across countries, but the number of households differs in inverse proportion to productivity: one country can have higher productivity but proportionately lower household count.

Revenue (\( R \)) for a Cournot firm in country \( i \) and selling in country \( j \) is given by the price in \( j \) times quantity of the firm’s sales. Price is a function of all firms’ sales.

8With the CES-CD version described in footnote 2, the Engels curve and the ratio \( Y/X \) is a smooth concave curve rather than the kinked version in Figure 1. Perfect aggregation never holds in the CES-CD case, and the aggregate demand for \( X \) will always be higher with a wider income distribution.
Assume zero trade costs and segmented markets, with \( c \) and \( F \) having the same value across countries. Then any firm that operates will sell in both markets and will sell the same in each market as any other firm regardless of that particular firm’s home country: \( X_{ij} = X_{ji}, \ X_{ij} = X_{ij} \). \( n \) will then denote the total number of firms from both countries activity in equilibrium, where \( n \) is determined by the usual free-entry zero-profit conditions. This zero–profit condition for a firm from \( i \) is as follows, with a corresponding condition for a firm from \( j \).

\[
p_i X_{ii} + p_j X_{ij} - c(X_{ii} + X_{ij}) - F = 0 \quad \text{zero profits} \quad (12)
\]

The result that gives the Cournot markup of a firm is fairly well known and I will not re-derive it here: the markup in \( j \) (defined on price \( p_j \)) is given by the firm’s market share divided by the price elasticity of demand. With zero trade costs and equal marginal costs, each firm in selling in a market has the same market share as any other firm regardless of home country. So the market share is always just \( 1/n \) and the markup is \( 1/(n\epsilon_j) \).

\[
m_{ij} = m_{ij} = p_j \left[ 1 - \frac{1}{n\epsilon_j} \right] = c \quad \text{or} \quad p_j = \left[ \frac{n\epsilon_j}{n\epsilon_j - 1} \right] c \quad \text{Cournot pricing} \quad (13)
\]

Market clearing in country \( j \) is given from earlier results, similarly for \( i \).

\[
p_j n X_{ij} + p_j n X_{ij} = \beta (M_j - M_j^0) \quad M_j^0 = \frac{(1-\beta)}{\beta} p_x z L_j \quad \text{market clears} \quad (14)
\]

The elasticity of demand \( \epsilon \) is given from (8) above. Also use (13) to replace \( p \) in \( m^0 \).

\[
\epsilon_j = \frac{m_j}{m_j - m_j^0} \quad m_j^0 = p_j \frac{1-\beta}{\epsilon} z = \frac{n\epsilon_j}{n\epsilon_j - 1} \frac{1-\beta}{\beta} z c
\]

Using the second equation of (15), the first becomes

\[
\epsilon_j = \frac{m_j}{m_j - \frac{n\epsilon_j}{n\epsilon_j - 1} \frac{1-\beta}{\beta} z c}
\]

which in turn reduces to

\[
\left[ \frac{n\epsilon_j - 1}{n\epsilon_j} \epsilon_j - 1 \right] = \frac{1-\beta}{\beta} \frac{z c}{m_j} \quad \epsilon_j \text{ is decreasing in } m_j, n \text{ constant} \quad (17)
\]

Recall that the number of firms \( n \) is common between \( i \) and \( j \). We then get a “cross-section” result from (17): comparing two countries, the higher per-capita income country will
have a lower price elasticity, higher markup and higher price level.

Add the zero-profit conditions for the representative firms in \( i \) and \( j \) in (12) together. There are four Cournot pricing equations (13) for the two market supplies for the representative firm from each country. Multiply both sides of the four pricing equations (13) by the relevant outputs and set to zero (move c terms to the left-hand side). Add these four together and subtract them from (12). This will yield the condition that:

\[
\frac{p_j(X_{ij}^* + X_{ij}^*)}{n \epsilon_j} + \frac{p_i(X_{ij}^* + X_{ij}^*)}{n \epsilon_i} = 2F \tag{18}
\]

Use the market-clearing equations in (14) to substitute for the revenue terms in (18).

\[
\frac{\beta (M_j - M_j^0)}{n^2 \epsilon_j} + \frac{\beta (M_i - M_i^0)}{n^2 \epsilon_i} = 2F \tag{19}
\]

and replace \( \epsilon_i \) and \( \epsilon_j \) with the expressions in (15). (19) becomes

\[
\frac{\beta (M_j - M_j^0)^2}{n^2 M_j} + \frac{\beta (M_i - M_i^0)^2}{n^2 M_i} = 2F \tag{20}
\]

Solving for \( n \), we have

\[
n = \left[ \left( \frac{(M_j - M_j^0)^2}{M_j} + \frac{(M_i - M_i^0)^2}{M_i} \right) \left( \frac{\beta}{2F} \right) \right]^{\frac{1}{2}} \text{ and } \tag{21}
\]

\[
n \epsilon_j = \left[ \left( \frac{(M_j - M_j^0)^2}{M_j} + \frac{(M_i - M_i^0)^2}{M_i} \right) \left( \frac{\beta}{2F} \right) \right]^{\frac{1}{2}} \frac{M_j}{M_j - M_j^0} \tag{22}
\]

Consider two identical countries with equal aggregate incomes \( M = M_i = M_j \)

\[
n \epsilon_j = \left[ \left( \frac{2(M - M^0)^2}{M} \right) \left( \frac{\beta}{2F} \right) \right]^{\frac{1}{2}} \frac{M}{M - M^0} = \left( \frac{\beta}{F} \right)^{\frac{1}{2}} \left[ \frac{M}{M^{1/2}} \right] \left[ \frac{M}{M - M^0} \right] = \left( \frac{\beta M}{F} \right)^{\frac{1}{2}} \tag{24}
\]

which (when inverted) gives a simple formula for the common markup.

\[
\frac{1}{n \epsilon} = \left[ \frac{F}{\beta M} \right]^{\frac{1}{2}} = \text{markup} \tag{25}
\]

The markup falls with a growth in aggregate income due to the pro-competitive
effects of entry (Venables 1985, Markusen and Venables 1988). But the interesting thing about (25) is that non-homotheticity washes out. Holding aggregate income constant, increase per-capita income (increase productivity offset by fewer households). This lowers the price elasticity of demand but this is exactly offset in this special case by more entry. Recognizing that this last result is derived for identical economies only, we can thus suggest that non-homotheticity does not have a “time-series” effect on markups as per-capita income grows over time, but does show up in the “cross-section” comparison between countries with different per-capita incomes.

Results for this section are illustrated in Figure 5. The “cross-section” result is shown in the left-hand panel. The two-countries are identical at a value of 0.5 on the horizontal axis. Then capital is transferred to h and labor to f, maintaining aggregate-income equality moving to the right, resulting in a rising markup in country h. With the price of $Y$ equalized between countries, this means that the price index is greater in country h as shown in Figure 5. The results on prices and markups is consistent with those in Wong (2003), Hsieh and Klendow (2007), Hummels and Lugovskyy (2009), and Simonovska (2009). Qualitatively, the same result occurs if we maintain relative endowments identical, but raise the productivity and lower the absolute endowment of country h and vice versa for country f.

The “time-series” result is shown in the right-hand panel of Figure 5 where the two countries are identical. Productivity is rising along the horizontal axis and absolute endowment lowered to maintain identical and constant aggregate incomes. The higher per-capita income moving to the right leads to a shift in consumption to $X$ and to an increase in intra-industry trade, inter-industry trade being zero. Thus trade volume increases relative to aggregate income. The same result will of course hold under monopolistic competition (Markusen 1986, Bergstrand 1989). A consequence is that gravity equations should show trade rising with per-capita income, aggregate income held constant, a topic discussed in the final section of the paper.

4. Thoughts on calibration, estimation and gravity

I want to conclude by offering some thought on the implications of the analysis for calibrated modeling (applied general-equilibrium analysis) and econometric estimation. A common procedure in AGE modeling is to assume a homothetic functional form such as Cobb-Douglas for example, and then used observed expenditure shares in the data to calibrate the share parameters of the Cobb Douglas.

Refer back to Figure 2 and think of expenditure shares at point A as being 0.5. The common AGE calibration method then assumes that any expansions of the economy will be on the Engels curve through A and C. However, the marginal expenditure share per household used in these experiments is 0.667 ($\beta = 2/3$). Thus if some counterfactual leads to an increase in per-capita income, the actual Engels curve is that through A and B in Figure 2. Furthermore, if two countries are observed to have different initial budget shares, such as one country at A and one at B in Figure 2, then they will be calibrated as having different but homothetic preferences when in fact they might have identical but non-homothetic
preferences.\textsuperscript{9}

There is a rather simple procedure that GE modelers can use to recalibrate their models to identical but non-homothetic preferences of the type used here. First, their data can be used as a cross-section data set to estimate the type of Stone-Geary utility function used here: the Stone-Geary yields the familiar linear expenditure system, which is the approach used in Bowen, Leamer and Sveikauskas (1987) and Hunter and Markusen (1989). This could, for example, generate a common value of $\beta$ in our two-good case, or more generally a set of $\beta$s across consumption goods. Then in each country the $\varepsilon$ or $Z$s could be calibrated by using the observed consumption shares in the data with the estimated $\beta$s. In our two-good case, a rearrangement of (5) gives us

$$s_x = \frac{p_x X}{M} = \beta + \frac{(\beta - 1) p_x Z}{M} = \beta + (\beta - 1) s_z$$

where $s_x$ is the observed share of $X$ in expenditure and $s_z$ is the calibrated share of (unobserved) $Z$ in income using an econometric estimate of $\beta$ common across countries. In counterfactual experiments not involving a change in the number of households $Z$ is held constant while it is allowed to vary with the number of households when that occurs. This procedure then distinguishes between movements along Engels curves AC and AB in Figure 1. Failing to do so as in standard models will lead to a mis-prediction about the effects of world productivity growth or capital deepening.

There is a long history of fitting Heckscher-Ohlin theory (Leamer 1980, Maskus 1985, Bowen, Leamer and Sveikauskas 1987, Davis and Weinstein 1991, Hakura 2001, Trefler 1995). A much simplified description of this literature is that it starts with the relationship $E = X - C$, where $E$ is the net export vector, $X$ is the production vector and $C$ is the consumption vector. $E$ is then converted into the “factor-content of trade”. A simple procedure, for example, is to estimate a common technology matrix $[A]$ where $a_{ij}$ is good $j$’s use of factor $i$. Then the measured factor content of trade for a country is $[A]E$. The predicted factor content of the production vector is just the country’s endowment vector, $V$.

A considerable amount of effort has gone into trying to improve the fit through modifications of the basic model but I think that it is reasonable to say that almost all of that effort has gone the production side; e.g., allowing for technique differences, price differences, and differences in factor quality. In most all cases I am aware of, $C$ is in fact not really fitted at all nor is it given by data: $C$ is assumed to be given by $sX^w$, where $s$ is the country’s share of world income and $X^w$ is the observed world production vector, equal to the world consumption vector. So the estimators impose the assumption of identical and homothetic preferences across countries and of course equal relative prices. The factor content of consumption is then $s[A]X^w = sV^w$. This gives the relationship $[A]E = (V - sV^w)$, where the left-hand side is the measured factor content of trade and the right-hand side is the predicted value. The typical result in the simplest formulations is that there is substantial

\textsuperscript{9}It is my understanding that the GTAP and GEMPAC models do allow the modeler to select non-homothetic preferences, but I am not familiar with their underlying structures.
“missing trade”: measured value is much less than the predicted value.

The findings of Hunter (1991), Cassing - Nishioka (2009) and Fieler (2010) should encourage researchers to devote some effort to improving the estimation of the consumption vector. Let \( C \) continue to denote the consumption vector and assume identical but non-homothetic preferences across countries and assume that commodity prices are the same everywhere (this is not a new assumption here, it is used in virtually all of the literature referred to). Let \( \beta_i \) denote the marginal share of good \( i \) in consumption and \( Z_i^k = z_i L^k \) be the fixed “endowment” of good \( z_i \) in country \( k \); \( \beta_i \) and \( z_i \) are identical across countries. The hypothesized demand for good \( C_i \) in country \( k \) when condition (4) holds is given by

\[
C_i^k = -Z_i^k + \beta_i (M^k + \sum_j p_j Z_j^k)/p_i
\]  

(18)

These demand values can be converted into shares of consumption \( \sigma_i^k \) by multiplying through by the price of good \( i \) and dividing by income

\[
\sigma_i^k = \frac{p_i C_i^k}{M^k} = \frac{\beta_i \left( \sum_j p_j Z_j^k \right) - p_i Z_i^k}{M^k}
\]

where

\[
\sum_i \sigma_i^k = \sum_i \beta_i = 1 \quad \text{and} \quad \sum_k \frac{M^k}{M^w} \sigma_i^k = \sum_k s^k \sigma_i^k = \sigma_i^w
\]

(19)\hspace{1cm}(20)

where superscript \( w \) denote world aggregate values. Note that this demand system satisfies adding up: summing (19) over \( i \) (goods), the final terms on the right-hand side sum to zero and the sum of the shares equals the sum over \( \beta \) which is one as noted in (20). The second equation of (20), which will be used shortly, is an implication that our demand system in (18) also satisfies global adding up ((18) applies to global consumption and income as well), implying that the share of \( i \) in world consumption \( \sigma_i^w \) is the income-share weighted sum of the consumption shares across countries \( (s^k = M^k/M^w) \).

Now introduce our assumption we have used throughout that \( Z_i^k = z_i L^k \) is strictly proportional to the labor supply \( L^k \) and that \( M^k = m^k L^k \), where \( m \) is average per-capita income.\(^{11}\) The \( L^k \) can be factored out of both the numerator and denominator of (19) and cancelled out. (19) can be written as:

\[
\frac{p_i C_i^k}{M^k} = \frac{\beta_i \left( \sum_j p_j Z_j^k \right) - p_i Z_i^k}{M^k}
\]

\[\text{where}\]

\[
\sum_i \sigma_i^k = \sum_i \beta_i = 1 \quad \text{and} \quad \sum_k \frac{M^k}{M^w} \sigma_i^k = \sum_k s^k \sigma_i^k = \sigma_i^w
\]

(19)\hspace{1cm}(20)

\(^{10}\)Equation (18) here is virtually identical to equation (7) in Bowen, Leamer, and Sveikauskas (1987). But they turn to questions quite different from the points made in the present paper.

\(^{11}\)Again, assume that all households have sufficient income to satisfy (4) and buy positive amounts of \( X \). Total income is just the average per-capita income times \( L \); summing over \( i \) households we get \( M = \sum (m_i/L)L = \overline{m}L \) where \( \overline{m} \) is average per-capita income, an observable.
\[
\sigma^k_i = \beta_i + \frac{\beta_i(\sum p_j z^k_j) - p_i z^k_i}{m^k} = \beta_i + \frac{\beta_i(\sum p_j z_j) - p_i z_i}{m^k} \quad z^k_i = z_i \forall k \quad (21)
\]

In our hypothesized preferences and demand system, the \( z_i \)'s are *common parameters* across countries and hence the country superscript can be dropped. The numerator of the right-hand term in (21) is identical across countries, and so (21) can be written as

\[
\sigma^k_i = \beta_i + \frac{\alpha_i}{m^k} \sum_i \frac{\alpha_i}{m^k} = 0 \quad (22)
\]

where the right-hand equation follow from (19).\(^{12}\) Since our demand system imposes identical preferences across countries, the \( \sigma_i^k \)'s only differ from world average values by differences across countries in per-capita income, which is easily observed and calculated. These \( \alpha \) and \( \beta \) parameters can be estimated across countries. Then let (22) give the fitted or predicted value of \( \sigma = \sigma^W \) and let \( \sigma^W_i \) denote the *observed* share of good \( i \) in world consumption, which in turn is equal to its share in world production \( X^w \). Let \( M^k \) denote the country's income and \( M^w \) denote world income (both scalars). Then the *predicted* demand for good \( i \) in our country is given by

\[
C^k_i = \sigma^k_i M^k = \frac{\sigma^k_i}{\sigma^W_i} \sigma^W_i M^k \quad (23)
\]

where

\[
\sigma^W_i M^k = \frac{X^w_i}{M^w} M^k = \frac{M^k}{M^w} X^w_i = s^k X^w_i \quad (24)
\]

Substituting (24) into (23), we have the predicted value of consumption.

\[
C^k_i = \left[ \frac{\sigma^k_i}{\sigma^W_i} \right] s^k X^w_i \quad C^k = s^k [\Sigma^k] X^w \quad \sum_k C^k = [I] X^w \quad (25)
\]

\(^{12}\)Not the similarity here to Deaton and Muellbauer’s (1980) “almost ideal demand system” (AIDS). Suppressing prices, this is \( \sigma^k_i = \beta_i + \alpha_i \log(m^k) \) for household \( k \) where \( \alpha > 0 \) is an income-elastic good instead of \( \alpha < 0 \) in my formulation. Without disputing that it is almost ideal for many applications, it does not allow aggregation and is not globally “regular”: it can lead to predicted shares greater than one and less than zero at extremes of income. Note that a negative predicted value of \( \sigma \) will occur in (22) if condition (4) does not hold. Thus both approaches need a lower bound on \( m \) and AIDS needs an upper bound as well. *Subject to (4)*, aggregation will hold in my approach, which allows \( m^k \) to be given by average income in (22).
where $\Sigma$ is a diagonal matrix with its $i$th diagonal element equal to $\sigma_i^k/\sigma_i^w$ and $I$ is the identity matrix. The equations in (25) satisfy adding up over countries, which follows from applying the right-hand equation of (20) to the second equation in (25). Let the predicted factor content of world consumption be given by

$$V^w = [A]X^w$$

so

$$s^k[A]X^w - s^kV^w = 0$$

(26)

Adding the right-hand equation of (26) to $(X - C)$, the predicted factor-content of trade is then

$$[A]E = [A](X^k - C^k) = [A](X^k - s^k\Sigma^kX^w)$$

$$= (V^k - s^kV^w) + s^k[A][I - \Sigma^k]X^w$$

(27)

where the last term on the right-hand side is the “correction” for non-homotheticity ($I$ is the identity matrix). Once again, adding up is satisfied when summing over countries $k$: the first bracketed term in (27) is zero by construction of $V^w$; the second term is zero as noted in the second and third equations of (25).

The equations in (27) give the intuition about how per-capita income differences due to different relative endowments combined with the assumption that labor intensity and low income elasticity are correlated tend to lower the predicted factor content of trade. Under these assumptions, there should be a systematic correlation such that export goods in which the country is relatively specialized in production ($X_i/sX_i^w > 1$ are goods in which the country is relatively specialized in consumption ($\sigma_i^k/\sigma_i^w > 1$) or the $i$th element of $[I - \Sigma^k]$ is $(1 - \sigma_i^k/\sigma_i^w) < 0$). Exports of good $X_i$ are then smaller under our assumptions. Import goods for which $(X_i/sX_i^w < 1$ are goods which have small shares in consumption, $(\sigma_i^k/\sigma_i^w < 1$ (or the $i$th element of $[I - \Sigma^k]$ is $(1 - \sigma_i^k/\sigma_i^w) > 0$). So net imports $E_i$ are smaller (less negative) under our assumptions.

Our assumptions then tend to reduce the predicted factor content of trade for any vector of world outputs. The difference between actual and fitted values of trade would then be reduced, reducing missing trade, given that trade volumes are over-predicted in the standard analysis where $(\sigma_i^k/\sigma_i^w) = 1$ by assumption. However, I emphasize that such an outcome is by no means trivial or guaranteed. If differences in per-capita incomes reflect differences in productivity rather than differences in capital / labor endowments, then predicted trade could be even larger than under homothetic preferences.

A final point has to do with standard gravity equations. A common practice is to put in separate terms for both aggregate income (or population) and per-capita incomes, yet there is rarely if ever justification for this. With homothetic preferences, aggregate income should soak up all the explanatory power. Yet estimates of this type invariably show an important, positive and independent role for per-capita income.

There are a great many formulations of the gravity equation in the literature. Let me
Giving one coefficient to the product of the two incomes is a somewhat unusual formulation. A more typical formulation is to give separate coefficients for both countries (e.g., Feenstra, Markusen and Rose (2001) use the relationship between $\alpha$ and $\beta$ to discriminate between alternative theories of trade. Give a quick and simple example provided by Frankel et. al. (1998). Let $M_i$ denote aggregate income, $P_i$ denote population for countries 1 and 2, and $T_{ij}$ the trade between countries i and j. The Frankel et. al. formulations is

$$\ln T_{ij} = \alpha \ln (M_1 * M_2) + \beta \ln \left( \frac{M_1}{P_1} \ast \frac{M_2}{P_2} \right) + \ldots. \quad (28)$$

In levels, this is equivalent to

$$T_{ij} = (M_1 * M_2)^\alpha \left( \frac{M_1}{P_1} \ast \frac{M_2}{P_2} \right)^\beta = (M_1 * M_2)^\alpha \ast \beta (P_1 * P_2)^{-\beta} \quad (29)$$

If the world is characterized by homothetic preferences, then aggregate income is all that matters, so the hypothesis is $\beta = 0$. If preferences are non-homothetic, then we should find $\beta$ to be non-zero, although in what direction is not obvious. If traded goods are income elastic (we should really include traded services here), then we should find $\beta > 0$. Alternatively, if differentiated goods are high income-elasticity goods then higher per-capita income countries will have a higher volume of intra-industry trade as discussed in section 3 and Figure 5. Frankel et. al.’s results are that $\alpha = 0.72$ and $\beta = 0.23$ with both coefficients significant at the one-percent level. The result that $\beta > 0$ supports the view that per-capita income is important and suggest that traded goods or especially differentiated goods are income elastic.

Parenthetically, the common result from gravity models that trade is per-capita-income elastic seems in contradiction with Trefler (1995), whose analysis points in the direction of high per-capita income countries trading less as a share of income (e.g., if income-elastic services are non-traded). Trefler (pp. 1038-1040) derives this in connection with his “endowment paradox”, which is that poor countries seem abundant in most factors and rich countries are scarce in most factors.

5. Summary

As suggested in the introduction, there are bits-and-pieces of theoretical and empirical analysis about the role or roles for per-capita income in determining trade flows. But there is little unity and by and large per-capita income is not given much of a place as an important determinant of trade. This paper tries to unify and connect the bits, and to offer some further ideas about how per-capita income might matter. I offer a “generic” model that I am hoping might prove to be useful for graduate teaching, a sort of all-in-one model that nests a number of other contributions.

13Giving one coefficient to the product of the two incomes is a somewhat unusual formulation. A more typical formulation is to give separate coefficients for both countries (e.g., $\ln T_{ij} = \alpha_i \ln M_1 + \alpha_j \ln M_2$ $\ldots$) with homogeneity implying that the $\alpha$’s sum to one. Feenstra, Markusen and Rose (2001) use the relationship between $\alpha_i$ and $\alpha_j$ to discriminate between alternative theories of trade.
The model imposes a variant of Stone-Geary preferences (used before by a number of authors) on top of a traditional 2x2x2 Heckscher-Ohlin model. Maintained hypotheses are that labor endowments in the HO model are proportional to the number of households and that the capital-intensive good in the HO model is the high income-elasticity-of-demand good. The latter assumption is testable and falsifiable. Results from the model offer a strictly demand-side explanation for a range of phenomena including (a) home bias in consumption, (b) the mystery of the missing trade, (c) a growing wage gap in an environment of growing productivity and (d) a role for the intra-country distribution of income similar to that found in the product-quality literature (higher inequality - more demand for luxury goods).

I then add an assumption of increasing returns to scale in the capital-intensive, high-income-elasticity industry with free entry and exit of firms, Cournot pricing and segmented markets: a common framework in the so-called new trade theory and strategic trade-policy literatures. This generates some interesting and testable results, in particular higher markups and higher price levels in higher per-capita-income countries, and more trade between higher per-capita-income countries, aggregate income held constant. As in the case of the competitive examples, some of the implications have already received good empirical support.

In both competitive and imperfect-competition cases, the effects of growth are quite different depending on whether it is growth in productivity or in neutral factor accumulation. This is potentially quite important in forecasting forward using econometric or CGE models: who would have predicted ten years ago that China would now be the world’s largest car market?

The paper concludes with a couple of suggestions about how the model might be useful in further calibration and estimation research and for including and interpreting per-capita income coefficients in gravity models. Models and econometric estimates based on homothetic demand and used for forecasting can significantly mis-predict the effects of growth through productivity and technical change.
REFERENCES


Table 1: Income elasticities of demand, budget shares for various consumption goods and services

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low income</td>
<td>high income</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Household furniture</td>
<td>0.76</td>
<td>food 0.50 &gt; 0.17</td>
</tr>
<tr>
<td>Fuel and power</td>
<td>0.81</td>
<td>furnishing 0.05 &lt; 0.09</td>
</tr>
<tr>
<td>Education</td>
<td>0.87</td>
<td>medical 0.04 &lt; 0.09</td>
</tr>
<tr>
<td>Clothing and footwear</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Beverages and tobacco</td>
<td>1.23</td>
<td>agricultural 0.09 &gt; 0.02</td>
</tr>
<tr>
<td>Other</td>
<td>1.25</td>
<td>food prod 0.15 &gt; 0.07</td>
</tr>
<tr>
<td>Recreation</td>
<td>1.42</td>
<td>bus serv 0.14 &lt; 0.23</td>
</tr>
<tr>
<td>Transport, communication</td>
<td>1.72</td>
<td></td>
</tr>
<tr>
<td>Gross rent</td>
<td>1.74</td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>1.91</td>
<td></td>
</tr>
</tbody>
</table>

Hunter and Markusen (1989) homogenize demand (shares same everywhere) trade flows increase by 29%

Low income elas <=> labor intensive

Hunter's (1991) counter-factual
Figure 1: Preferences and their properties

Engel's curve: $Oy_0A$

$m_0$: minimum income for consuming $X$

![Engel's curve diagram](image-url)
Figure 2: Growth through productivity versus factor accumulation
Figure 3: Wage gap from the demand side (identical countries)

- Productivity level: countries identical
- Both lines flat with homothetic demand
- Relative price of X, r/w ratio

Relative price of X, r/w ratio

Productivity level: countries identical
Figure 4: Home bias and the mystery of the missing trade

Correlation between production / consumption shares = home bias?

Country h's share of world capital (h's labor share = 1 - capital share)
Figure 5: Per-capita income, markups, price levels, trade volume
Countries have identical and constant aggregate income

Markups and price level: relative endowment differences (or productivity differences)

Proportional markup difference: country h to f
Proportional price index difference: country h to f
All values = 0 with homothetic demand

Country h capital endowment, country f labor endowment (countries identical at 0.5)

Volume of trade related to per-capita income
(aggregate income held constant)

Volume of trade relative to aggregate income
All values = 1 with homothetic demand

Per-capita income level, countries identical