Institute for Agricultural Policy, Market Research and Economic Sociology, University of Bonn

Discussion Paper 2005:1

Two ways of estimating a transport model

Torbjörn Jansson

The series "Agricultural and Resource Economics, Discussion Paper" contains preliminary manuscripts which are not (yet) published in professional journals, but have been subjected to an internal review. Comments and criticisms are welcome and should be sent to the author(s) directly. All citations need to be cleared with the corresponding author or the editor.

Editor: Thomas Heckelei
Institute for Agricultural Policy, Market Research
and Economic Sociology, University of Bonn
Nußallee 21
53115 Bonn, Germany

Phone: +49-228-732332
Fax: +49-228-734693
E-mail: heckelei@agp.uni-bonn.dede

Two ways of estimating a transport model

Torbjörn Jansson

Abstract

In this article, it is shown how the parameters of a transport model can be estimated in a way that, in contrast to previously used methods, utilizes observations of regional prices as well as of trade costs. The proposed method uses bi-level programming to minimize a weighted least squares' criterion under the restriction that the estimated parameters satisfy the Kuhn-Tucker conditions for an optimal solution of the transport model. We use Monte-Carlo simulations to trace out some properties of the estimator and compare it with a traditional calibration method. The analysis shows that the proposed technique estimates prices as well as trade costs more efficiently.

Keywords: Spatial equilibrium, transport model, bi-level programming
JEL-classification: C15, F11

1 Introduction

The transport model treated in this article is a common component of spatial price equilibrium models. It has been analyzed in several central articles in linear programming, for example in Koopman's original article from 1947, Enke's ingenious "solution by electric analogue" (1951) and Samuelson's formalized treatment (1952). In Dantzig's work on linear programming (1966), the transport problem is referred to as "The classical transport problem." This paper is not concerned with the solution of the transport model, which has been thoroughly treated for more than fifty years, but turns instead to the empirical specification of the model. In fact, during the long history of this established problem, little attention has been paid to the estimation of the input data for the problem.

The transport cost minimization problem can be written as

$$\begin{aligned} \min_x \quad & \sum_{ij} c_{ij} x_{ij} & & \\ \text{s.t.} \quad & e_i + \sum_j (x_{ji} - x_{ij}) = 0 & & [p_i] \\ & x_{ij} \geq 0 & & [v_{ij}] \end{aligned} \qquad (1)$$

where c_{ij} is the trade cost from region i to region j, x_{ij} is the traded quantity and e_i is excess demand in i. Letters in square brackets after the restrictions symbolize the dual values of the constraints.

In this article, we consider the trade of a single homogeneous good, and assume that consistent data on regional excess demand is available. Furthermore, we assume that observations of trade costs between regions as well as regional prices are available, but associated with measurement errors. Observed prices are likely to be inconsistent with observed trade costs and excess demand under the assumption that they constitute an equilibrium solution to model 1.

Spatial price equilibrium models frequently contain a similar transport cost minimization model as component. Examples range from the early publications of Judge and Wallace (1958) and Takayama and Judge (1964) to the more recent contributions of Litzenberg, McCarl and Polito (1982), Peeters (1990) and Guijardo and Elizondo (2003)—to name just a few.

In the cases known to this author, including the publications just cited, there was either no calibration at all, or the models were calibrated by

1. solving the trade cost minimization problem using the observed trade costs,
2. taking the dual values of the market clearing restrictions p_i as prices and
3. shifting the prices so that some important price is matched precisely.

Step 3 is possible because the first order conditions only contain pair-wise *price differences*. Indeed, one of the market clearing restrictions is redundant, because we know that for a solution to the transportation problem to exist, the sum of all regional net demand must be zero (Dantzig 1966), implying that if there are k markets, then if k-1 of them clear, all of them must clear. Because only price differences are identified, one *numerator price* can be chosen arbitrarily and the remaining prices are determined by those price differences.

Obviously, this method for determining regional prices for a transport model does not use any direct observations of regional prices except for the numerator price. The remaining regional price information is extracted from trade costs and excess demand. This procedure is henceforth referred to as "traditional" and abbreviated TRAD.

In this paper we (1) demonstrate an alternative method, a bi-level estimation program (BLEP), for calibrating the input data for a transport model that uses observations of regional prices, and (2) show that BLEP estimates regional prices more efficiently than TRAD and that this increased efficiency in estimating prices does not come at the expense of a less efficient estimation of trade costs.

The rest of this article is outlined as follows: In the next section, the BLEP is presented in some detail, and it is given a geometric interpretation. Then we use a three-region example model to deduct hypotheses about the behavior of the two estimators. The hypotheses are analyzed using Monte-Carlo simulations, where the performance of the two estimators is evaluated using generated data. The results of the simulations are analyzed and compared to the hypotheses formed. A final section summarizes and discusses the results.

2 The bi-level estimation program

2.1 The estimator

Heckelei and Wolff (2003) propose estimating parameters of agricultural supply models by using optimality conditions as estimating equations. Jansson and Heckelei (2004) show how a similar technique can be applied to the estimation of a transport model, where a large number of inequalities renders the estimation numerically difficult. The current paper contributes to this strand of research by calibrating the parameters of a transport model by direct estimation of the optimality conditions of problem 1, using a least squares objective, and analyzing the finite sample properties of the estimator. The resulting optimization problem, given by equations 2-7 belongs to the class *bi-level programming problems* (BLPP). The name is due to the fact that it is one programming problem, the estimation, that has another programming problem, in this case the transport problem represented by its optimality conditions, in the constraints.

$$\min_{p,c} \quad \sum_{ij} \left(c_{ij} - c_{ij}^{obs}\right)^2 + \sum_{i} \left(p_i - p_i^{obs}\right)^2 \tag{2}$$

$$\text{s.t.} \quad e_i + \sum_{j} \left(x_{ji} - x_{ij}\right) = 0 \tag{3}$$

$$c_{ij} - p_j + p_i = v_{ij} \tag{4}$$

$$x_{ij} v_{ij} = 0 \tag{5}$$

$$c_{ij} = c_{ji} \tag{6}$$

$$x_{ij} \geq 0, \; v_{ij} \geq 0 \tag{7}$$

The general BLPP is difficult to solve, so a few words about solution techniques are appropriate, although a substantial treatment of that subject is beyond the scope of this article. Several different solution methods were tested, including approximation by smooth reformulations as suggested by Facchinei, Jiang and Qi (1999), a branch-and-reduce algorithm called BARON implemented as solver for the modelling language GAMS and the method proposed by Jansson and Heckelei. Those three methods all produced similar results and were therefore assumed capable of solving the problem satisfactorily. The latter method was finally chosen because it runs faster than the others, thus enabling a larger number of simulation experiments.

2.2 A geometric interpretation

It may be worthwhile to dwell a moment on the BLEP to obtain a better understanding of what the estimator does. It turns out that, in the case where the criterion function to be minimized is the sum of squared deviations and the model to be estimated is a linear model, it has an intuitive geometric interpretation. To be specific, we consider the following simple BLEP, where we estimate a linear programming problem in one variable y that depends on a parameter x, and restrictions as follows:

$$
\begin{aligned}
\min_{x,y}\ & (x - x_o)^2 + (y - y_o)^2 \\
\text{s.t.}\ & \min_{y|x}\ y \\
& \text{s.t.} \begin{array}{rcr} -y - x & \leq & -3 \\ -y + x & \leq & 2 \\ y - x & \leq & 2 \\ y + x & \leq & 8 \end{array}
\end{aligned}
\tag{8}
$$

x_o and y_o are observations, and we want to pick x and y that minimize the upper level objective and where y solves the inner problem treating x as given. We note that (x,y) that minimize $(x - x_o)^2 + (y - y_o)^2$ also minimize $\sqrt{[(x - x_o)^2 + (y - y_o)^2]}$, which is the *Euclidean metric*, i.e. the distance, between the estimated point (x,y) and the observation (x_o,y_o).

In figure 1 the restrictions of problem (8) are drawn as lines, the observed point (x_o,y_o) as a plus sign, and level curves for the criterion function as concentric circles around the observation. All points on a circle have the same distance from the plus sign and hence the same objective values in the criterion function. Following Bard (1998), we call the area enclosed by the restrictions (where the circles are not dashed) the *constraint region S* of the bi-level programming problem.

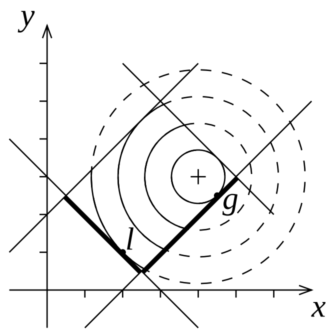

Figure 1. A simple BLEP with OLS criterion and linear inner problem.

The projection of S onto the x-axis is denoted by $S(X)$, and is a convex subset of X with the property that for each $x \in X$ at least one, but possibly several, solutions to the LP exists. If we form the set of all pairs (x,y) where x is in $S(X)$ and y solves the LP, we have the so-called *inducible region*. It is marked with heavy lines in the figure. We seek the point in the inducible region that is closest to the observation.

When the inner problem is an LP, the inducible region is a piecewise linear equality constraint derived from the faces of S (Bard 1998). In the general case, it is non-convex, so there may be several local optima. In figure 1, there is a local optimum at the point l and the global optimum is found at g. The non-convexity of the inducible region causes difficulties for the solution of the problem, and is one important reason that special solution methods frequently are needed for BLEPs.

3 Hypotheses about the estimators

It is desirable that an estimator on average is close to the true parameter. We call this *efficiency* and measure it by the *mean squared error* (MSE) (Greene 2003). MSE is the mean squared deviation of an estimate from the true parameter value. Efficiency is a relative measure, so what we would like to know is if one of the estimators is more efficient than the other. To our aid, we use the fact that MSE can be split into a variance and a bias component using

$$MSE[\hat{\theta} \mid \theta] = E[(\hat{\theta} - \theta)^2] = Var[\hat{\theta}] + (Bias[\hat{\theta} \mid \theta])^2 \tag{9}$$

where θ is the true parameter value and $\hat{\theta}$ the estimator. In this section, we formulate hypotheses about the efficiency of the estimators, to be tested later. We proceed by qualitative reasoning, based on the structure of the two methods, to deduct hypotheses about variances and biases. If one estimator turns out to be less biased as well as having less variance than the other, we conclude that it is more efficient. If on the other hand one estimator is less biased but has higher variance than the other, the qualitative reasoning in this section does not allow us to say that one estimator is more efficient than the other. To make the reasoning easier to follow, we first present the results of this qualitative section. The following two sections in this article report setup and results of simulation experiments designed to analyse each of the properties.

1. BLEP is a more efficient estimator of regional prices than TRAD, because the BLEP estimates have both less variance and less bias than the TRAD estimates.

2. We cannot *a priori* say that either estimator is a more efficient estimator of trade costs. On the one hand, BLEP estimates have a bias that TRAD

estimates lack, but on the other hand the variances of the BLEP estimates are lower. The simulation experiments reported below suggest that this hypothesis can be strengthened.

3. The variances of the price estimates are heterogeneous, in other words the variance is different in different regions. It is more heterogeneous if estimated with TRAD than with BLEP.

4. The variance of the cost estimates is heterogeneous when estimated with BLEP but not when estimated with TRAD.

The reasoning is illustrated in a three-region, single good model. We assume that there exist true parameter values that represent an equilibrium solution to this model. We can think of TRAD and BLEP in a mechanical way as two devices that accept observations of regional prices and costs as inputs and produce estimates of prices and costs as outputs, or rather, as a one-one relation from the price-cost space into itself (onto the inducible region). The principle procedure followed here is (i) to note that if regional prices and trade costs are observed *without* errors, then both BLEP and TRAD return the same equilibrium solution, i.e. the true parameter values, and (ii) to introduce "known measurement errors" systematically for certain price and trade cost positions in order to see how the equilibrium solutions returned by TRAD and BLEP change in response. Or, if the estimator is thought of as a mapping, (i) to note that the true parameter values is a point in the inducible region that maps to itself and (ii) to see how images of points symmetrically dispersed around the true values are dispersed around the image of the true values.

The Kuhn-Tucker conditions for an optimal solution of (1) imply that if a trade flow is connecting two regions, the price difference must be precisely the trade cost. Increasing the trade cost between the regions will firmly push the equilibrium prices apart, and decreasing it will pull them together. However, the complementarity restrictions will truncate the response of prices to changed observed trade costs, because trade always flows along the cheapest path. If the trade cost between two regions that trade with one another is continuously increased, then at some point the trade route will not longer be the cheapest one, so trade will take another path. Further increases of that trade cost will have no effect on prices. This fact is the cause of the biases of the estimators, and it is exemplified below.

Figure 2 shows three regions A, B and C, with B being a net importing region and A and C net exporters. The left and right panels of the figure show two of the three possible trade flows that would clear all markets. To be specific, let us assume that the true regional prices are $p_A = 100$, $p_B = 109$ and $p_C, = 104$ and the trade costs $c_{AB} = 9$, $c_{AC} = 5$ and $c_{CB} = 5$ and symmetric as in equation (6). In this case, trade will flow as in the left hand panel.

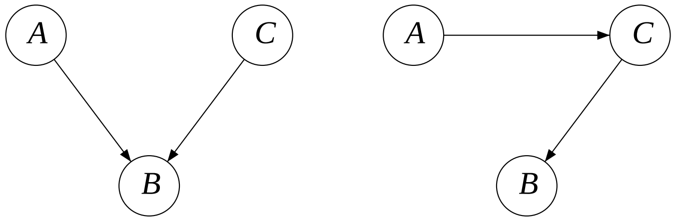

Figure 2. Three region model with A and C net exporters and B a net importer. Two possible market clearing solutions.

To start with, we feed the true parameter values into TRAD and BLEP, and as this observation is consistent with equilibrium, both methods map this observation to itself, i.e. both methods return precisely the observation. Now, what happens to the estimates if upon observation a random variable from a symmetric distribution with mean of zero is added to the trade cost c_{AB} and all other observations remain undisturbed? The symmetric errors will have a *biasing* effect on prices, regardless if they are estimated with TRAD or BLEP, as illustrated by the following numerical example:

Example: *We measure prices and costs of the model in figure 2 twice, and after each measurement we use the observation to estimate the true parameters with TRAD and BLEP. Only* c_{AC} *is measured with errors, all other trade costs and prices are observed with their true values. The observations of* c_{AC} *are*

(i) $c_{AC} = 10$

(ii) $c_{AC} = 0$

TRAD. (i) The trade cost minimizing solution is the same as that without the error, so trade will still flow as in the left panel of the figure. The dual values of the markets with the numerator price p_A added will equal the true prices, because the flow *AC* is still not used. The costs will, as always with TRAD, be the observed ones: $c_{AB} = 9$, $c_{AC} = 10$ and $c_{CB} = 5$. (ii) It is cheaper to transport via *ACB* than via *AB*, so trade will divert from *AB* to *ACB* as in the right panel of figure 2, the prices will be $p_A = 100$, $p_B = 105$ and $p_C = 100$, and costs $c_{AB} = 9$, $c_{AC} = 0$ and $c_{CB} = 5$. Conclusion: In this case, only negative errors that are larger than 1.0 influence the price estimates, because the second cheapest trade route is 1.0 unit more expensive than the cheapest one. The price estimates for *B* should systematically turn out lower than the true prices in this setup, as would the price in *C*.

BLEP. (i) The observation is a point in the inducible region, so the estimator will accept the observation unaltered and will measure a deviation of zero. In the estimated model, trade will flow as in the left panel of the figure. Nothing will happen to the prices because the flow *AC* is still not used, and the estimated costs will be $c_{AB} = 9$, $c_{AC} = 10$ and $c_{CB} = 5$ as with TRAD. (ii) The observation is not in the inducible region, so the estimator will look for the closest point of the inducible region using the least squares criterion. The best solution means using the trade flow *ACB* and not *AB*, choosing the prices $p_A = 101.9$, $p_B = 108.2$ and $p_C = 102.9$, and the trade costs $c_{AB} = 9$, $c_{AC} = 0.952$ and $c_{CB} = 5.381$. Again, only the negative error with absolute amount greater than 1.0 influences the estimation. As with TRAD, a symmetric measurement error with a mean of zero causes the estimated prices to deviate from the true values in only one direction (positive for p_A and negative for p_B and p_C), i.e. being estimated with bias, but less biased than with TRAD.

The example can be modified to consider measurement errors on other trade costs than c_{AC}. For example, a reasoning similar to that above suggests that symmetric measurement errors on c_{CB} would cause the price in *C* to be *overestimated* on average and that of *B* to be underestimated. So, errors on c_{AC} leads to the price in *C* being *underestimated* and errors on c_{CB} work in the opposite direction. Which effect will be stronger? Perhaps there is a way to compensate for those biases in the estimator to obtain unbiased estimates? The biases seem to depend on excess demand, which is known, but also on true prices and costs, which are assumed to be unknown. At this point, therefore, we are satisfied with concluding that the estimates will be biased, that they are more biased for TRAD than BLEP, but that we cannot formulate a general rule for the size and direction of the bias.

From this simple example we learn that symmetrically distributed errors on trade cost observations have a biasing effect on price estimates because positive and negative errors influence the estimator differently. The cause of this effect is of course the Kuhn-Tucker conditions, allowing the price difference between two regions to be smaller than but never larger than the trade cost—the inequalities built into the Kuhn-Tucker conditions truncate certain errors away.

Prices only occur in equation 4, and there always pair-wise. Hence, it is not primarily the estimated *prices* that are biased, but the price *differences*—which equal the trade costs for the used trade flows. So, in the example we see that BLEP also tends to bias the trade cost estimates: from two observations symmetrically distributed around the true value (5.0) the estimator delivers the two estimates 10.0 and 0.952. The average estimate is then 5.476, so there is a bias of $5.476 - 5.0 = 0.476$. The bias resulting from this source would always be positive, i.e. towards an overestimation of the cost. However, the estimator will consider the errors on all other measured parameters of the model, which may influence the bias in positive or negative direction. As there are not infinitely many prices and trade costs in the model, the biases will most probably not cancel each other out, but it is difficult to determine the direction of the bias *a priori*.

Above, we let the observations deviate from the true values for one trade cost at the time and observed how the estimated point deviated from the true values. A similar reasoning can be applied to measurement errors on prices. Assuming that all parameters happen to be observed at their true values except for one price that is observed with error yields the following conclusions for the two estimators: TRAD will not contend with such price errors at all unless it is the numerator price that is subject to the error, because all other price observations are discarded. However, a measurement error on the numerator price will be added to all other prices as well. BLEP will distribute the price error over several regions, and also over the estimated trade costs. It is difficult to predict the signs and sizes of the estimated errors on costs and prices resulting from a measurement error for one price item, as it depends on the interconnections via trade flows of the region with other regions. Other *prices* are obviously influenced with the same sign as the error. For *trade cost* estimates it can have positive, negative or zero biasing effect depending on trade relations.

Clearly, the *variance of each trade cost* estimated with TRAD will be precisely the same as that of the measurement error of that trade cost. But what happens with BLEP? Here there seem to be two mechanisms working in opposite directions: On the one hand, the measurement errors of prices will influence the trade cost estimates, tending to increase the variances. On the other hand, the least squares estimator will attempt to distribute deviations over as many parameters as possible thus decreasing the variance. An educated guess worth testing is that the second effect dominates, i.e. the variances of the trade cost estimates with BLEP are generally lower than those of TRAD.

The *variance of each price* estimated with TRAD will generally be *higher* than that of prices estimated with BLEP. The "estimated" numerator price will be precisely the observed numerator price, so for that price, the variance of the estimator will be that of the observations. Because all other prices are tied to the numerator price by trade costs, the estimates of those other prices will receive the error from the enumerator price but also from each used trade link (each with a trade cost measured with error!) that separates them from the region with the numerator price. As a contrast, BLEP does not maintain any numerator price, but distributes errors over *all* prices and trade costs. Thus, the variances of the prices estimated with BLEP should be lower than those estimated with TRAD. Furthermore, this mechanism will cause the variance of the price estimates to be different in different regions if estimated with TRAD, following the general rule that the further away a region is from the numerator region in terms of used trade links, the higher the variance of the price estimate will be.

The variance of regional price estimates should be different in different regions, *heterogeneous*, if estimated with *either* method. In *any* region that is a net importer, say *B*, the equilibrium price is precisely equal to that of at least one neighboring region, say *C*, from which *B* imports the good, plus the trade cost. We call this the *supply price* of *C* in *B*. Note that some authors, e.g. Anderson and

Wincoop (2004) use the term supply price differently. The Kuhn-Tucker conditions imply that the supply price in *B* of any regions not delivering to *B* is higher than p_B. This is the case for *A* in the right panel of figure 2. If we measure the trade cost from *C* to *B* too high, i.e. with a positive error, the estimated price in *B* will be higher than the true price, but not higher than the supply price of *A* in *B*. Now, if the supply price of *A* in *B* is, by chance, *close* to the supply price of *C* in *B*, and trade can divert into the situation of the left panel of figure 2, not much will happen to the price in *B* if c_{CB} is measured higher than the true value. A part of the error is "truncated away." If, conversely, the supply prices are *far apart*, regional prices are more strongly linked to the trade costs, and a measurement error of the trade cost from *C* to *B* will have greater influence on prices.

The heterogeneity of the variance of estimated prices also spills over into the variance of trade costs estimated with BLEP, as this method distributes errors between prices and trade costs. To conclude, both methods are biased estimators of the variances of prices, and BLEP is also a biased estimator the variances of trade costs.

4 Simulation experiments

The small sample properties of the estimation are analyzed using simulation techniques. The basic idea is to generate *m* randomly chosen "true models," and then estimate each model *n* times (the simulation size is *n*), each time adding errors to the true prices and costs. We thus obtain *m* samples of *n* observations of estimated trade cost matrices and price vectors. Throughout this paper we use $m = 100$ and $n = 500$.

The *m* models, each with ten regions, are generated by drawing regional excess demand from the uniform distribution $[-10,10]$ and trade costs from the uniform distribution $[20,100]$. The excess demand of one region is set to the negative of the sum of excess demand in all other regions to make the problem feasible. The transport model (1) is solved, and the dual values of the market balances plus a constant of are 120 taken as true regional prices. In the following, the index denoting the model to which a certain price or trade cost belongs is omitted.

Each of the *m* models is estimated *n* times with TRAD and BLEP, each time with errors added to all true prices and trade costs. The errors are sampled from the normal distribution with mean of zero and standard deviation 6. With this standard deviation, the rule of thumb "plus or minus three standard deviations" lets us expect that the major part of the errors are in the interval $[-18,18]$. By construction, true trade costs are in $[20,100]$, so with a numerator price of 120 the smallest possible true trade cost as well as price is 20. Hence, adding an error of standard deviation 6 and mean zero will rarely result in negative observed values. Still, they may occur, and to prevent that, the sampled errors are truncated to lie within the interval [−19,19]. The errors are truncated upwards as well as downwards to avoid truncation being a source of biases.

In the next section, we address the hypotheses put forward in the previous section by analysing MSE, variances and biases of prices and trade costs estimated with TRAD and BLEP. Since equation (9) holds for each parameter in each model, we can compute the mean of each term over all prices or costs in each model, obtaining mean MSE (MMSE), mean squared bias (MSBIAS) and mean variance (MVAR), for which it holds that MMSE = MSBIAS + MVAR. The means are computed in order to obtain an overview over the large number of parameters estimated in the simulation exercise.

5 Results

This section presents the results of the simulation experiments in relation to the hypotheses previously formed. It is subdivided into three parts: (i) efficiency of price estimates, (ii) efficiency of trade cost estimates and (iii) heterogeneity of variances.

5.3 Efficiency of price estimates

Result 1 *BLEP is a more efficient estimator of regional prices than TRAD.*

The simulation experiments confirm the hypothesis that BLEP is a more efficient estimator of regional prices than TRAD. Figure 3 shows MMSE for price estimates in all models. Each point is the average MSE for all regional prices in one model.

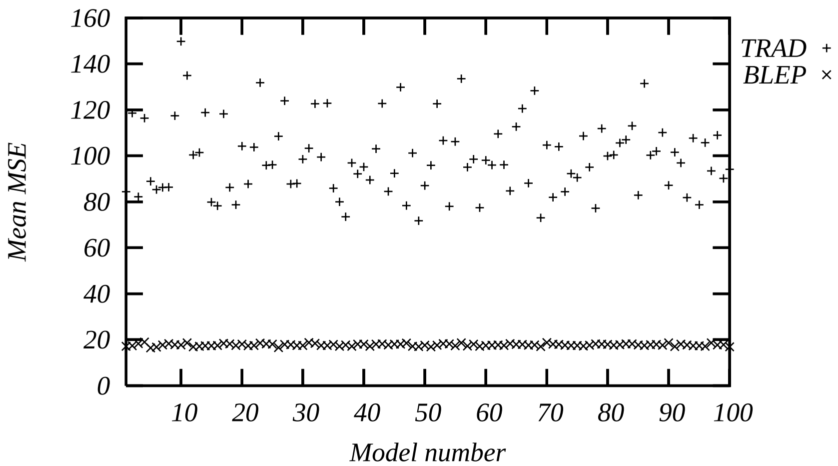

Figure 3. Mean MSE for price estimates for each model.

The results show that BLEP not only delivers more efficient estimates of prices, but also that the efficiency is stable across different models. In other words, it does not depend upon the true data constellation. In contrast, TRAD is less efficient in all cases, and additionally, the efficiency seems to depend on the data constellation. As we will see, the greater efficiency of BLEP regarding price estimates is attributable to less bias as well as less variance, as the qualitative reasoning above suggests.

Result 2 *Both BLEP and TRAD estimate prices with bias, but the bias is smaller for BLEP.*

Figure 4 shows the mean squared bias of price estimates in all models. MSBIAS of prices estimated with TRAD fluctuate strongly between models, whereas the biases of prices estimated with BLEP are much more stable and also smaller. Most of the large biases come from the TRAD estimator.

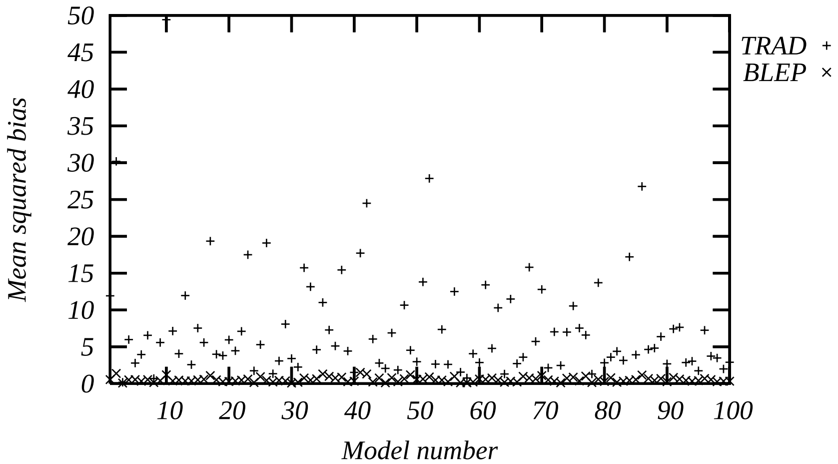

Figure 4. Mean squared bias of price estimates of all models.

It may also be of interest to analyse the biases of the individual regional price estimates (not the mean squared bias). Table 1 shows descriptive statistics of the sample of estimation biases of regional prices. Neither the average nor the median of the biases is far from zero, indicating that there are about as many positive biases as there are negative ones. The larger variance of the biases of TRAD supports the hypothesis that TRAD generally produces price estimates with larger biases. The larger biases also appears in the line "SABIAS", which is the sum of absolute biases. SABIAS of TRAD is more than three times that of BLEP.

Table 1: Descriptive Statistics of Biases of Price Estimates

	TRAD	BLEP
mean	-0.320	0.006
variance	7.500	0.544
median	-0.118	-0.004
SABIAS	1929.660	567.130

According to the reasoning in the previous section, we would expect TRAD to systematically estimate biased prices in some regions in some models. One way of testing this is to perform two-sided t-tests of the hypothesis "the average price estimated for an arbitrary region in an arbitrary model does not equal the true price," with the null hypothesis that they are equal. The test statistic is computed as

$$T_i = \left| \frac{\sqrt{n}\left(p_i - \overline{p}_i\right)}{s_i} \right| \quad \text{with} \quad s_i = \frac{\sum_{l=1}^{n} \left(\hat{p}_i^l - \overline{p}_i\right)^2}{n-1} \tag{10}$$

where $\hat{p}_i^l$ is the estimated price in region i for estimation l with $l \in \{1,\ldots,n\}$, $\overline{p}_i$ the average of the n estimations of the price in region i, and s_i the sample standard error of the price estimates in region i and in each model (model index still omitted!). We compare this with the upper critical values of Student's t-distribution.

The result is that for TRAD, the null hypothesis is rejected in 567 cases (of 1000 possible, 100 models with 10 prices per model) at the 95 percent level and in 504 cases at the 99 percent level. For BLEP, the null hypothesis is rejected in 569 cases at 95 percent level and 475 cases at 99 percent level. The test seems to support the hypothesis that both estimators are biased, but does not make any clear distinction between them.

The power of the t-test is probably low, as the price estimates are likely to violate normality, because of the "truncating mechanism" making the distributions asymmetric. Still, there is an indication that price estimates are systematically biased, and that they are more biased when estimated with TRAD than BLEP. So, TRAD is likely to deliver price estimates that on average lie further away from the true prices than does BLEP.

Result 3 *The variance of prices estimated with TRAD is greater than that of prices estimated with BLEP.*

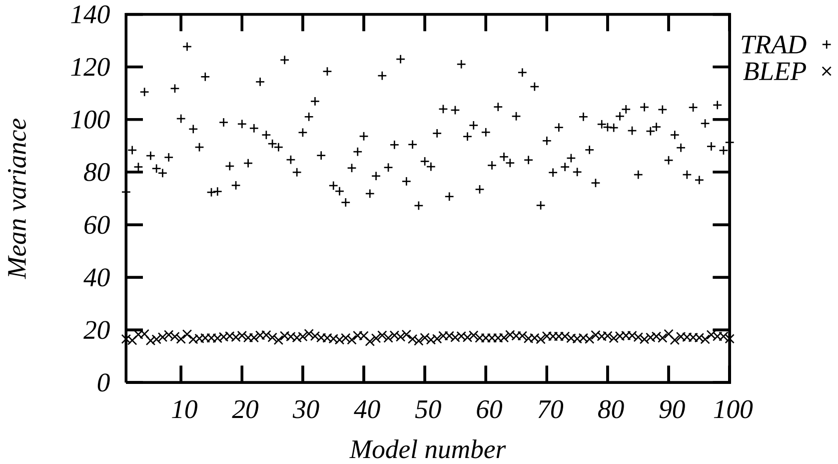

Figure 5. Sample variances of price estimates, pooled together for each model.

Figure 5 shows the pooled sample variance of price estimates in each model estimated with TRAD and BLEP. If $r = \{1,\ldots,m\}$ indexes the models, the pooled sample variance[1] s_r^2 of the prices of model r is computed as $s_r^2 = \frac{1}{k}\sum_{i=1}^{k}\left(s_{ri}^2\right)$, with s_{ri}^2 indicating the squared sample standard deviation of price i in model r as defined above, and k indicating the number of regions. As can be seen in the figure, TRAD estimates generally have a much higher variance. The variances of prices estimated with TRAD seem to depend more strongly upon the underlying true model than is the case for BLEP. The highest pooled sample variance of the TRAD estimates is about twice the lowest one, whereas the variances of the BLEP estimates are closer together.

It is noteworthy that in all models, the pooled variance for BLEP is clearly smaller than the variance used in the sampling process (36), whereas it is much bigger for TRAD. The pooled sample variances reflect the behavior of the underlying non-pooled variances, described in more detail below under the hypothesis about heterogeneity. Obviously, no statistical test is necessary to see that the variance of the TRAD price estimates is greater than that of the BLEP.

[1]Because each price is estimated the same number of times, the pooled variance turns out to be the plain average MVAR.

5.4 Efficiency of trade cost estimates

Result 4 *BLEP is a more efficient estimator of trade costs than TRAD.*

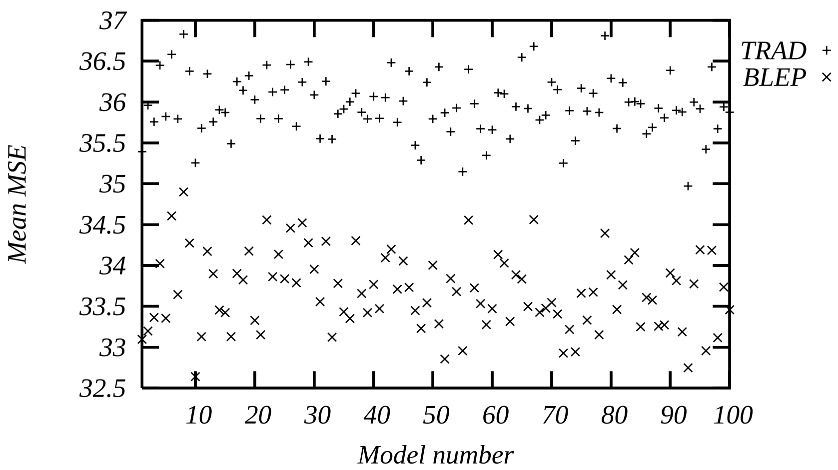

Figure 6. Mean MSE for trade cost estimates for each model.

Figure 6 shows the mean MSE for all trade cost estimates in each model. MMSE for BLEP is lower than for TRAD in all models, but the differences are not as obvious as for the price estimates (figure 3). It also looks as if the efficiency of BLEP is somewhat more sensitive to different data constellations than TRAD, because the BLEP points appear to be vertically more dispersed. Below, MMSE of trade cost estimates is split up into bias and variance components.

Result 5 *BLEP but not TRAD produces biased trade cost estimates.*

The TRAD trade cost estimates cannot, by construction, be systematically biased. They are simply the unaltered observations, so their being biased would mean that there is something wrong with our data generating process. However, we have only a finite sample, so the sample mean (the mean of any estimated cost item taken over the *n* repetitions) may very well deviate from the true trade cost.

For BLEP, the qualitative discussion above suggests that the inequalities could cause the trade costs to be systematically biased for some region pairs, but in an unpredictable direction. In figure 7, the MSBIAS of the trade cost estimates in all models are shown. All values are small, and it is not immediately clear whether BLEP is more biased than TRAD, but the tendency is certainly visible, because points further away from zero generally belong to BLEP.

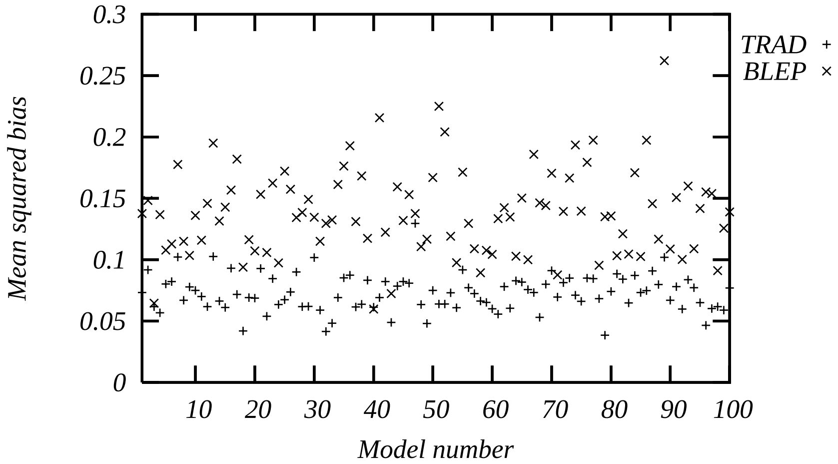

Figure 7. Mean squared bias of trade cost estimates in each model.

To further investigate the question whether the BLEP cost estimates actually are more biased than the TRAD estimates, we perform a t-test similar to the one performed for price biases above. As we do not know if a given trade cost will be over- or underestimated, we make the test two-sided. For TRAD, the number of rejections of the null hypothesis is close to the number that would be expected, namely 5.42 percent of the cases (244/4500) at the 95 percent level and 1.36 percent of the cases (61/4500) at the 99 percent level. For BLEP, the number of rejections is higher, with 721 rejections at the 95 percent and 403 at the 99 percent level, thus supporting the hypothesis.

As for price biases, the nature of the bias (truncation by the inequality) may tend to make the distribution of the estimates with BLEP (but not with TRAD!) asymmetric, so the sample will probably not be normally distributed, weakening the power of the t-test.

If all the computed trade cost biases are considered a sample, we get the sample statistics shown in table 2. The average bias is close to zero for both methods, and so are the variances and the medians of the biases. The sum of absolute biases of all trade costs in all models, SABIAS, is higher for BLEP than for TRAD, also supporting the result.

Table 2. Descriptive Statistics of Biases of Trade Cost Estimates

	TRAD	BLEP
mean	-0.007	0.041
variance	0.073	0.136
median	-0.007	0.013
SABIAS	964.690	1260.942

Result 6 *The variance of trade costs estimated with TRAD is greater than that of those estimated with BLEP.*

We expect the pooled sample variance of the estimated trade costs for TRAD to be precisely 36, which is the variance used in the data generation process. Furthermore, the hypothesis states that the pooled sample variance per model (see above) of the estimates performed with BLEP should be lower. Figure 8 shows the pooled sample variances of both methods for each of the 100 models. The data seems to support the hypothesis, because the TRAD data points are nicely dispersed around 36 and all BLEP data points lie below the lowest TRAD data point.

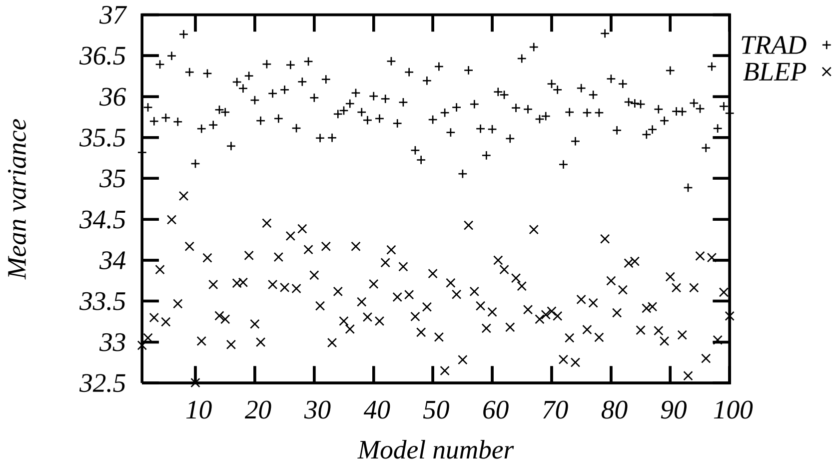

Figure 8. Mean variance of estimated trade costs, pooled together for each model.

If we look at the underlying data in the form of the non-pooled sample variances, the view is more differentiated. It seems that the sample variances of the cost estimates are more dispersed across trade links within each model with BLEP than with TRAD. This observation is further discussed in connection with the hypothesis regarding homogeneity of variances below (figure 10).

The conclusion is that the variances of trade cost estimates *in general* are lower with BLEP than TRAD, but that this conclusion must not hold for an *arbitrarily* selected trade cost estimate.

5.5 Heterogeneity of variances

Result 7 *The variance of the price estimates is heterogeneous, i.e. the variance is different in different regions. It is more heterogeneous if estimated with TRAD than with BLEP.*

A quick look at the data supports this result. Figure 9 shows the sample variance (not pooled) of the first 200 prices estimated, i.e. of all prices in the first 20 models. Each model has ten regions, so there is a total of two hundred points for each of TRAD and BLEP shown in the figure. It can be seen that the variances of the different regional price estimates fluctuate strongly between the TRAD estimates, whereas the variances seem much more homogeneous for BLEP. All BLEP points are in the thick band at the bottom of the plot. Above that band comes a row of plus signs, which is the TRAD estimates of the numerator prices, all of which have the variance 36 (the sampling variance). Above that row lie all the other TRAD estimates. The higher the variance of a price estimate, the more trade links are probably separating it from the numerator price.

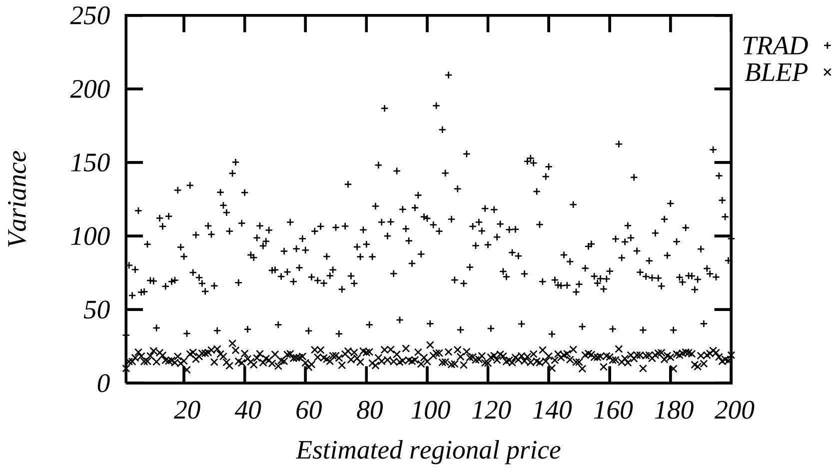

Figure 9. Sample variance of estimates of individual prices

The variances of the TRAD estimates are clearly heterogeneous. However, it is difficult to tell whether the variances of the BLEP estimates are homogeneous or not, that is if the fluctuations observed are random outcomes of the same distribution. If we would estimate the same price item another n times, would we then get a similar or different sample variance? We want to test the hypothesis "the variances of prices differ between at least two regions in the estimated model" with the null hypothesis "the variances are equal in all regions of the model." To do this, a *Bartlett's* test (see NIST/SEMATECH) is performed for each model m. The results indicate that in 100 models out of 100, TRAD has produced heterogeneous estimates at the 95 percent significance level, whereas BLEP seem to have done so in 99 cases. At the 99 percent level there is no change, i.e. the null hypothesis is still rejected in 100 out of 100 cases for TRAD, and in 99 of 100 cases for BLEP.

However, the Bartlett's test is sensitive to deviations from normality, and we know that the price estimates are biased. Hence, the results may be due to a skewed distribution, not to heterogeneity. To double-check, we perform also a Levene's test (ibid.) for heterogeneity, a test that is less sensitive to deviations from normality. The test can be performed using deviations from mean or from the median. Both were tried, with similar results. The following results are for tests with the mean. The test statistics indicate that price estimates are biased (at the 95 percent level) in all 100 models with both TRAD and BLEP. At the 99 percent level the results are similar, with 99 rejections for BLEP and 100 for TRAD. So, it seems like the price estimates are likely to be heterogeneous with both methods, albeit the visual impression from figure 9 clearly is that the problem is smaller for BLEP than for TRAD.

Result 8 *The variance of the cost estimates is heterogeneous when estimated with BLEP but not when estimated with TRAD.*

Figure 10 shows the sample variance of the first 200 trade cost estimates with TRAD and BLEP. The first impression is that there is less difference between the methods than was the case for the price estimates. The variances of the TRAD estimates are, as expected due to the data generation method, dispersed around 36. The variances of the BLEP estimates seem to be generally smaller, as previously discussed, and more dispersed, supporting the hypothesis. A lot of the points in the figure coincide. These are trade costs for trade links that are not used regardless of cost, so the value need not be modified in order to reach consistency.

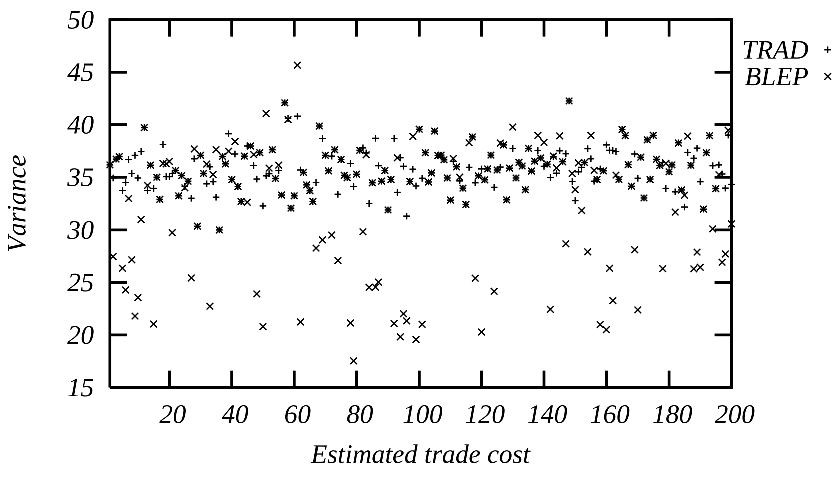

Figure 10. Variance of estimates of individual trade costs.

The tests for heterogeneity detect clear differences between the TRAD and the BLEP estimates: The Bartlett's statistic for the hypothesis "the variances of all cost estimates in each model are not equal," with the null hypothesis "all variances in each model are equal" fully supports the hypothesis. The null hypothesis is rejected at the 95 percent significance level in only four models of 100 with TRAD, and in no model at the 99 percent level. For BLEP, the null hypothesis is rejected in all 100 models on both the 95 percent and the 99 percent level.

To double check, the Levene's test was performed also for trade cost estimates with results similar to those of the Bartlett's test. The null hypothesis is rejected in six of 100 models for TRAD at the 95 percent level and in three cases at the 99 percent level. For BLEP, the Levene's test rejects the hypothesis in all models at both levels of significance. The hypothesis thus seems to be firmly corroborated.

6 Discussion

We conclude that BLEP is a more efficient estimator than TRAD of prices as well as trade costs. For prices, BLEP estimates have smaller biases as well as smaller variances than TRAD. For trade costs, the BLEP estimates are biased whereas the TRAD estimates are not. However, the biases of the BLEP estimates are more than compensated for by lower average variances. Variances of trade cost estimates are heterogeneous if estimated with BLEP but not with TRAD. In other words, the variances of prices are estimated with bias in both methods.

The BLEP performs better than TRAD in almost all disciplines. Are there no drawbacks? Clearly, one drawback is that BLPPs in general are difficult to solve. However, with increasing computing capacity and the development of new solver software, that argument is rapidly losing its strength. And for the incumbent problem—the transport model—existing techniques seem to be able to handle the difficulties.

References

Anderson, J. E. and E. van Wincoop (2004). Trade costs. *Journal of Economic Literature 42*, 691–751.

Bard, J. (1998). *Practical bilevel optimization*. Kluwer academic publishers.

Dantzig, G. (1966). *Lineare Programmierung und Erweiterungen*. Springer-Verlag Berlin.

Enke, S. (1951). Equilibrium among spatially separated markets: Solution by electric analogue. *Econometrica 19*, 40–47.

Facchinei, F., H. Jiang, and L. Qi (1999). A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming 85*, 107–134.

Greene, W. H. (2003). *Econometric analysis* (Fifth ed.). Prentice Hall.

Guajardo, R. G. and H. A. Elizondo (2003). North american tomato market: a spatial equilibrium perspective. *Applied Economics 35*, 315–322.

Heckelei, T. and H. Wolff (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximim entropy. *European Review of Agricultural Economics 30*, 27–50.

Jansson, T. and T. Heckelei (2004). Estimation of a transport model using a mathematical equilibrium constraints. Technical Report 1501, Global Trade Analysis Project.

Judge, G. G. and T. D. Wallace (1958). Estimation of spatial price equilibrium models. *Journal of Farm Economics 15*, 801–820.

Koopmans, T. C. (1947). Optimum utilization of the transportation system. *Econometrica 17*, Supplement.

Litzenberg, K., B. A. McCarl, and J. Polito (1982). A decomposition solution for spatial equilibrium problems. *American Journal of Agricultural Economics 64*, 590–594.

NIST/SEMATECH (2004, Dec). NIST/SEMATECH e-handbook of statistical methods. http://www.itl.nist.gov/div898/handbook/, National Institute of Standards and Technology.

Peeters, L. (1990). A spatial equilibrium model of the EU feed grain sector. *European Review of Agricultural Economics 17*, 801–820.

Samuelson, P. A. (1952). Spatial price equilibrium and linear programming. *American Economic Review 42*, 283–303.

Takayama, T. and G. G. Judge (1964). Spatial equilibrium and quadratic programming. *Journal of Farm Economics 46*, 67–93.