Southern Economic Journal 1998, 65(1), 169–177

Experimental Results on Expressed Certainty and Hypothetical Bias in Contingent Valuation

Karen Blumenschein,* Magnus Johannesson,† Glenn C. Blomquist,‡ Bengt Liljas,§ and Richard M. O'Conor†

Use of the contingent valuation method is controversial among economists because it is based on hypothetical rather than real choices. Previous experiments have suggested that the commonly used dichotomous choice contingent valuation method leads to hypothetical bias, i.e., overestimates the real willingness to pay. We carried out an experiment to compare the dichotomous choice contingent valuation method with real purchase decisions for a consumer good. We confirm previous findings that hypothetical yes responses overestimate real purchase decisions, but we cannot reject the null hypothesis that definitely sure yes responses correspond to real purchase decisions.

1. Introduction

In order to determine the optimal level of nonmarket goods, such as environmental quality, information about the willingness to pay for the good is needed. The contingent valuation method has been developed to measure the willingness to pay for environmental changes (Mitchell and Carson 1989). The method means that individuals are asked about their hypothetical willingness to pay for a defined good. The use of the contingent valuation method is controversial among economists (Diamond and Hausman 1994; Hanemann 1994). The controversy is centered around the extent to which the hypothetical choices in the contingent valuation method correspond to real economic choices. Despite this controversy, relatively little experimental work has been carried out directly comparing hypothetical and real willingness to pay.

A number of different techniques exist to elicit the willingness to pay in a contingent valuation study. Currently, the most commonly used elicitation format is the dichotomous choice approach (Hanemann 1994). In a dichotomous choice contingent valuation question, a subject responds yes or no to a hypothetical question about paying a specified price for a defined good. By varying the price in different subsamples of respondents, the mean willingness to pay for the good can be estimated (Hanemann 1994).

Cummings, Harrison, and Rutström (1995) recently presented data from an experiment

* College of Pharmacy, University of Kentucky, Lexington, KY 40536-0082, USA.

† Department of Economics, Stockholm School of Economics, Box 6501, S-113 83 Stockholm, Sweden; E-mail hemj@hhs.se; corresponding author.

‡ Department of Economics and Martin School of Public Policy and Administration, University of Kentucky, Lexington, KY 40506-0034, USA.

§ Department of Economics, Lund University, Box 7082, S-220 07 Lund, Sweden.

We are grateful to Glenn Harrison, Jason Shogren, and two anonymous referees for helpful comments. An earlier draft was presented at the Southern Economic Association meetings held November 21–23, 1997, in Atlanta, Georgia.

Received July 1997; accepted March 1998.

comparing the dichotomous choice contingent valuation method for measuring willingness to pay with real decisions. Three experiments were carried out on three different consumer goods (an electric juicemaker, a calculator, and a box of chocolates). For each good, the proportion of hypothetical yes responses exceeded the proportion of real yes responses. In early work by Bishop and Heberlein (1979) and in more recent work by Nape et al. (1995), dichotomous choice contingent valuation questions to measure willingness to accept also have been shown to lead to similar problems of overestimation. Overestimation in hypothetical willingness to pay questions has also been shown for hypothetical referenda (Bjornstad, Cummings, and Osborne 1997; Cummings et al. 1997) and open-ended willingness to pay questions (Neill et al. 1994; Loomis et al. 1996).

The overestimation of the dichotomous choice contingent valuation approach noted by Cummings et al. (1997) was also confirmed in another experiment on private goods by Johannesson, Liljas, and Johansson (1998). In that experiment, the hypothesis that a more conservative interpretation of the approach, where only definitely sure yes responses are counted as yes responses, correctly predicts real purchase decisions was also tested. Definitely sure yes responses were found to significantly underestimate the real yes responses and thus provided a lower bound for the real willingness to pay. A related approach was also used in a recent study by Champ et al. (1997), which compared hypothetical dichotomous choice questions about donating a specified amount to a public good with actual donations to the public good. They also assessed the certainty of the hypothetical donation responses on a 1–10 scale from very uncertain to very certain. They found that hypothetical donations significantly exceeded real donations but that there was no significant difference if only subjects that were very certain of their yes responses were counted as real yes responses.

In this study, we carried out another experiment to compare the dichotomous choice contingent valuation method with real decisions for a consumer good. We confirm the results of Cummings, Harrison, and Rutström (1995) that hypothetical yes responses overestimate real purchase decisions, but we cannot reject the null hypothesis that definitely sure yes responses correspond to real purchase decisions. The design of the experiment is outlined below along with the results and some concluding remarks.

2. Design of the Experiment

Our experiment involved a pair of sunglasses made by UVEX. These sunglasses are used as protective eyewear in professional laboratories in the U.S. and are not typically available in ordinary retail shops. The sunglasses are similar in style to sunglasses used for sporting activities such as bicycling and playing volleyball. We ordered the sunglasses directly from the wholesaler at a bulk price of $9 per pair. Undergraduate students taking an introductory economics class at the University of Kentucky College of Business and Economics were included in the experiment. Participation in the experiment was voluntary and each student received $5 for participating.[1] In total, 133 subjects participated in the experiment, which was carried out in August 1996.

[1] Giving the students $5 may have led to some endowment effect. However, for this to affect the conclusions of the experiment, the endowment effect has to affect the hypothetical and real purchases in different ways. Given the small number of purchases in the experiment, the endowment effect does not seem of great actual concern.

The subjects were divided into two groups. The first group received a hypothetical dichotomous choice question followed by a real purchase question, and the second group received only the real purchase question. The hypothetical question was worded in the following way:[2]

> We are interested in how much you are willing to pay for the sunglasses displayed at the front of the room. These sunglasses are made by UVEX and block 99% of ultraviolet light (both UVA and UVB). They have an antifog and antiscratch coating as well as sideshields, which make them suitable for driving, hiking, and other outdoor activities. These sunglasses have moveable temples so that they can be adjusted for individual fit. The sunglasses will be brought around for you to examine before you tell us your answer.
>
> Assume that you were offered the chance to buy these sunglasses here and now with your own money. Observe that you are not actually being offered the chance to buy the sunglasses but that you are asked to think about what you would do if you were offered the chance to purchase the sunglasses here and now. Would you buy the sunglasses here and now at a price of $5.00? Note that you are to answer if you would buy the sunglasses here and now at a price of $5.00 and not if you would buy the sunglasses on any other occasion at this price. Please circle yes or no below when you have made up your mind.

This question was similar to the dichotomous choice questions used by Cummings, Harrison, and Rutström (1995), with the exception that we placed much greater emphasis on the fact that we were interested in whether they would purchase the good here and now rather than on any other occasion. This study was thus not an attempt at an exact replication of the Cummings, Harrison, and Rutström (1995) experiment. The "here and now" aspect of the decision was stressed because it is possible that the difference between hypothetical and real responses found in the Cummings, Harrison, and Rutström (1995) study was due to differences in how respondents perceived the hypothetical versus the real questions. As noted by Cummings, Harrison, and Rutström (1995, p. 261), subjects might have viewed the hypothetical question as "would you ever pay $X for this good?" and the real question as "would you pay $X for this good right now?"

The hypothetical dichotomous choice question was followed by a question in which the subjects who answered yes were asked if they were "probably sure" or "definitely sure" about buying the sunglasses here and now at a price of $5.[3] In what we call "a conservative interpretation of the dichotomous choice contingent valuation method," only definitely sure yes responses were considered to be yes responses and the probably sure yes responses were considered to be no responses. The rationale for this conservative interpretation is that, if the dichotomous choice contingent valuation method leads to a systematic overestimation of the real yes responses, as indicated in the Cummings, Harrison, and Rutström (1995) and the Johannesson, Liljas, and Johansson (1998) studies, the yes responses with the method can be divided into two categories: (i) yes answers where the individual would actually buy the good in a real situation and (ii) yes responses where the individual would not buy the good in a real situation. To correctly predict real choices, it is thus necessary to sort out the real yes responses from the false yes responses (Blackburn, Harrison, and Rutström 1994). The hypothesis we wanted to test was that the real yes responses corresponded to the definitely sure yes responses and that the false yes responses corresponded to the probably sure yes responses.

---

[2] All data and questionnaires are available from the authors on request.

[3] The subjects who answered no to the dichotomous choice contingent valuation question received a follow-up question, where they were asked if they were "probably sure" or "definitely sure" that they would not buy the sunglasses here and now at a price of $5. This information was not used in the analysis since there is no evidence that a hypothetical no answer does not correspond to a real no answer. Of the 55 subjects that answered no to the hypothetical question, 44 were definitely sure that they would not buy the sunglasses here and now and 11 were probably sure that they would not buy the sunglasses here and now.

After the question about the certainty of the yes/no answer, the subjects were offered the chance to actually buy the sunglasses. This question was worded in the following way:

> We would now like to give you the opportunity to actually buy the sunglasses here and now for $5.00. Note that you do not have to give the same answer to this question as to the previous hypothetical question. If you answer yes, you will have to pay $5.00 at the end of the experimental session and you will then receive the sunglasses. Would you like to buy the sunglasses for $5.00? Please circle yes or no below when you have made up your mind.

Since it cannot be excluded that the real purchase decision is affected by the fact that the subjects have already received the hypothetical question, a second group of subjects received only the real question. This question was almost identical to the real question in the within-subjects group.[4] In contrast to Cummings, Harrison, and Rutström (1995), we used more than one price in the experiment to test that the results were consistent with a downward sloping demand curve. One group of subjects received a price of $1 and a second group received a price of $5.[5]

We tested whether the proportion of hypothetical yes responses differed significantly from the proportion of real yes responses. The proportion of hypothetical yes responses was analyzed both according to the standard interpretation of the yes/no responses and the more conservative interpretation based on the definitely sure yes responses. Our statistical tests involved two types of comparisons. In the between-group comparisons, the hypothetical responses for the part of the sample given both treatments were compared with the real responses for the part of the sample given only the real treatment. In the within-group comparisons, the difference between the hypothetical and real responses was examined using only that part of the sample subjected to both treatments. A nonparametric contingency table chi-square test for differences in proportions was used to test whether the proportion of hypothetical and real yes responses differed in the between-subjects comparisons (Newbold 1991). For the within-subjects comparisons (the subjects that received both a hypothetical and a real question), we used the nonparametric sign test to test whether the proportion of hypothetical and real yes responses differed (Newbold 1991). In addition to comparing real and hypothetical yes responses, we also compared the real yes responses between the two samples to test the null hypothesis of no ordering effect (i.e., that the real yes response is not affected by the fact that the subject has already received a hypothetical question). The nonparametric contingency table chi-square test for differences in proportions were used for this test (Newbold 1991). All the nonparametric tests were carried out at each of the two price levels.

We also evaluated our between-subjects hypotheses using regression analysis, controlling for the background characteristics of the individuals and the price level. We tested whether a hypothetical/real dummy variable was statistically significant in a probit regression equation (Greene 1993). The background characteristics collected in the study were sex (1 = male), age, how many pairs of sunglasses the subject owned (pairs), and the monthly spending money over and above room and board (money). The proportion of men in the sample was 0.45, the mean age was 20.02 years (SD 2.27), the mean monthly spending money was $217.99 (SD 199.74), and the subjects on average owned 1.49 pairs of sunglasses (SD 0.99). We did not evaluate our

---

[4] The only difference was that the word "now" was deleted from the first sentence and the second sentence was deleted. The question also started with the first paragraph of the hypothetical dichotomous choice question.

[5] Nape et al. (1995) also used more than one price in their comparison of hypothetical and real willingness to accept, and Johannesson, Liljas, and Johansson (1998) used more than one price in their comparison of hypothetical and real willingness to pay.

Table 1. Experimental Results, Number (%) of Yes Responses

	Within Subjects			
	Hypothetical			Between Subjects,
	Standard	Conservative	Real	Real
$1	9/33 (27)	2/33 (6)	2/33 (6)*	2/31 (6)*
$5	1/32 (3)	1/32 (3)	1/32 (3)	1/37 (3)

* Significantly different from the standard hypothetical yes responses at the 5% level.

within-subject hypotheses with probit regression analysis since this would create a correlation in the error terms for the two data points for the same individual.

3. Results

Table 1 shows the number and percentages of yes responses in the different groups. The percentage of yes responses decreased at the higher price in all the groups, consistent with a downward sloping demand curve. At the low price ($1), the percentage of standard hypothetical yes responses exceeded the percentage of real yes responses, and the null hypothesis of no difference was rejected at the 5% level for both the within-sample and the between-sample comparisons. At the high price ($5), the percentage of standard hypothetical yes responses was the same as the percentage of real yes responses, and the null hypothesis could not be rejected at the 10% level for the within-sample comparison or the between-sample comparison.[6] The percentage of conservative hypothetical yes responses was the same as the proportion of real yes responses at both the low price and the high price, and the null hypothesis of no difference could thus not be rejected at the 10% level for the within-sample comparison or the between-sample comparison.[7] The percentage of real yes responses was the same in both groups at both the low and the high price, and the null hypothesis of no ordering effect could not be rejected at any of the prices.

The probit regression results with a hypothetical/real dummy variable are shown in Table 2.[8] In the first regression equation in Table 2, the null hypothesis tested was that there is no difference between the standard hypothetical yes responses and the real yes responses in the between-subjects comparison. Since the hypothetical dummy variable was significant at the 5% level, the null hypothesis was rejected. The price variable was significant with a negative sign

[6] Since only one individual out of thirty-two wanted to hypothetically buy the good at the $5 price, it was, however, impossible for the hypothetical yes responses to significantly exceed the real yes responses at this price. Therefore, the $5 results added very little information to the study outside of demonstrating a downward sloping demand curve.

[7] Apart from comparing the overall proportion of conservative yes responses with the proportion of real yes responses, we can also test whether the individuals in the "hypothetical group" (that received the hypothetical question first followed by the real question) made real purchase decisions consistent with the conservative interpretation of the yes responses. Of the 55 people that said no to the hypothetical question, no one answered yes to the real question. Of the eight people who answered yes and were probably sure, one answered yes to the real question and, of the three people who answered yes and were definitely sure, two answered yes to the real question. The conservative interpretation of the dichotomous choice contingent valuation responses thus correctly predicted 63 out of the 65 real responses in the hypothetical group.

[8] Due to nonresponse for one subject on the age and spending money questions (in the hypothetical group; the subject answered no to the standard hypothetical question at a price of $5), the number of observations was 132 in the regression analysis.

Table 2. Results of Probit Estimations with Dummy Variable for Experimental Group (t-Ratios in Parentheses)

	Standard Hypothetical Versus Real Responses	Conservative Hypothetical Versus Real Responses	Real Responses[a]
Constant	−3.39**	−3.93***	−4.02***
	(−2.50)	(−2.68)	(−2.61)
Price	−0.32***	−0.15	−0.16
	(−2.73)	(−1.27)	(−1.30)
Sex	−0.96**	−0.70	−0.92
	(−2.22)	(−1.34)	(−1.62)
Age	0.13**	0.14**	0.16**
	(2.20)	(2.21)	(2.33)
Money	−0.00054	−0.00086	−0.0041
	(−0.45)	(−0.53)	(−1.46)
Pairs	−0.043	−0.019	0.15
	(−0.22)	(−0.08)	(0.62)
Hypothetical	0.91**	0.21	0.28
	(2.22)	(0.46)	(0.59)
N	132	132	132
Log-likelihood	−31.03	−20.08	−18.60
LRI[b]	0.27	0.18	0.24
Chi-square	22.89***	8.65	11.61*

[a] This equation contains a comparison of the real responses in the two groups (the variable "hypothetical" is thus included to test for an ordering effect rather than for a difference between hypothetical and real questions).
[b] LRI, likelihood ratio index.
*, **, *** Significant at the 10, 5 and 1 % levels, respectively.

at the 5% level, consistent with a downward sloping demand curve. Sex and age were the only significant background characteristics. The probability of a yes response was lower for men than women and increased with age. The regression equation was significant at the 1% level according to the chi-square value. With the explanatory variables held equal to their mean values, the predicted probability of a yes response at the $1 price level was 4% for the real treatment and 21% for the hypothetical treatment.

In the second regression equation in Table 2, the null hypothesis tested was that there is no difference between the conservative hypothetical yes responses and the real yes responses in the between-subjects comparison. Since the hypothetical dummy variable was not significant at the 10% level, the null hypothesis could not be rejected. The only significant variable in this regression equation was age, and the probability of a yes response increased with age. The regression equation was not significant at the 10% level according to the chi-square value.

In the third regression equation in Table 2, the null hypothesis of no ordering effect between the real yes responses in the two samples was tested. Since the group dummy variable was not significant at the 10% level, the null hypothesis could not be rejected. As in the second regression equation, the only significant variable was age, and the probability of a real yes response increased with age. The regression equation was significant at the 10% level according to the chi-square value. A reason for the low significance of the second and third regression equations is probably that so few respondents actually purchased the good. The proportion of yes responses was only 4.5% in the second and third regression equations in Table 2, which means that there was little variation to explain. The inability of these regression equations to explain the variation

in the dependent variable also suggests that there may be considerable random noise in the experimental data.

We also tested for structural differences between the regressions for the different experimental groups (Nape et al. 1995). This was done by adding interaction terms between the hypothetical group dummy variable and all other variables. A likelihood ratio test was then carried out to test whether this unrestricted model differed from the restricted model without the sample dummy variable and the interaction term. The likelihood ratio test statistic for this test was 17.76 (6 d.f.) in the comparison of standard hypothetical yes responses and real yes responses (significant at the 1% level) and 10.54 in the comparison of conservative hypothetical yes responses and real yes responses (not significant at the 10% level). For the comparison of the real yes responses between the two experimental groups, the likelihood ratio test statistic was 7.82 (not significant at the 10% level). These tests for structural differences were thus consistent with the above tests of the significance of a hypothetical/real dummy variable. The regression results were also consistent with the results of the nonparametric tests.

4. Concluding Remarks

The lack of market data makes it necessary to rely on other information to estimate the willingness to pay for nonmarket goods such as environmental quality. The contingent valuation method has therefore been developed to measure the willingness to pay for environmental changes. Since the contingent valuation method is based on hypothetical rather than real choices, its use is controversial among economists. To carry out experiments comparing real and hypothetical willingness to pay is one way of assessing the validity of the contingent valuation method. The experiment reported in this paper provides some information about the correspondence between hypothetical and real willingness to pay for the most commonly used contingent valuation elicitation technique, the dichotomous choice contingent valuation approach.

In the experiment, one group of subjects was asked if they would hypothetically purchase a pair of sunglasses here and now at a specified price and another group of subjects was actually offered the opportunity to purchase the sunglasses at the same price. The proportion of subjects that said they would hypothetically purchase the sunglasses significantly exceeded the number of subjects that actually purchased the sunglasses. This overestimation of the real purchases in the hypothetical group is consistent with the results of the Cummings, Harrison, and Rutström (1995) and Johannesson, Liljas, and Johansson (1998) studies. Our study thus provides additional evidence of the presence of hypothetical bias in dichotomous choice contingent valuation questions.

In this experiment, we also tested the hypothesis that a more conservative interpretation of the dichotomous choice contingent valuation approach, where only definitely sure yes responses are counted as yes responses, correctly predicts real purchase decisions. We could not reject the null hypothesis that the definitely sure yes responses corresponded to the real yes responses. This finding differs from the result in the Johannesson, Liljas, and Johansson (1998) study, where the hypothetical definitely sure yes responses underestimated the real yes responses. This difference in results could be due to a number of factors. The sample size (133 in this study vs. 242 in the Johannesson, Liljas, and Johansson [1998] study) and thus the statistical power to reject the null hypotheses differed between the studies. The studies were also carried out in different countries (Sweden and the U.S.) and languages, and the follow-up

question to distinguish between probably and definitely sure yes responses may have been interpreted in different ways in the two countries. There were also some differences in the design of the experiments that may have affected the results.[9]

It is also interesting to compare our results concerning the correspondence between hypothetical definitely sure yes responses and real yes responses with the results of the recent study by Champ et al. (1997). They found that hypothetical donations significantly exceeded real donations but that there was no significant difference if only subjects that were very certain of their yes responses were counted as real yes responses. Thus, the results are similar to our findings; however, there are some important differences between the studies that limit the comparability between them. We compared real and hypothetical questions about willingness to pay for a private good. Champ et al. (1997), on the other hand, compared real and hypothetical questions about voluntary donations to a public good, where real donations were interpreted as a lower bound on the willingness to pay for the public good. The question about the certainty of the hypothetical yes responses, used to sort out the respondents that were certain of their yes responses, also differed between the studies. We used a dichotomous question about the certainty, whereas Champ et al. (1997) used a continuous scale.

A major weakness with treating only definitely sure yes responses as yes responses is that there is little theory to support this approach. Without understanding the reasons why individuals might answer yes to a hypothetical question and then follow that with a no to a similarly posed real question, it is hard to understand why only those who respond definitely yes in follow-up questions would be those who later respond yes to a real question. In order to better understand why hypothetical and real contingent valuation questions may differ, it is important to analyze the motivation in individuals' answers to hypothetical questions and to develop a theory of how individuals respond to hypothetical questions. Due to the lack of a theory of mistakes in nonarbitraged markets that could be used for treating definitely sure yes responses as real yes responses, our result on the correspondence between the two has to be treated with great caution. Our results, together with the recent results of Champ et al. (1997), however, suggest that the certainty of a hypothetical yes response may be an important predictor of a real yes response. Further work on the possibility of calibrating hypothetical willingness to pay responses based on the degree of certainty of the responses may thus be worthwhile.

References

Bishop, Richard C., and Thomas A. Heberlein. 1979. Measuring values of extramarket goods: Are indirect measures biased? *American Journal of Agricultural Economics* 61:926–30.

Bjornstad, David, Ronald Cummings, and Laura Osborne. 1997. A learning design for reducing hypothetical bias in the contingent valuation method. *Environmental and Resource Economics* 10:207–21.

Blackburn, McKinley, Glenn W. Harrison, and Elisabet E. Rutström. 1994. Statistical bias functions and informative hypothetical surveys. *American Journal of Agricultural Economics* 76:1084–8.

[9] As compensation to the students for participation in the study, we used cash payments, whereas Johannesson, Liljas, and Johansson (1998) used commodity payments (a voucher for a cinema ticket). In the Johannesson, Liljas, and Johansson (1998) experiment, an introductory paragraph was included in the experimental instructions explaining that the experiment was part of a research project to test methods to value goods for which there are no markets. No such introductory statement was made in this experiment. In the Johannesson, Liljas, and Johansson (1998) study, subjects who wanted to buy the good were also given the opportunity to take an interest-free loan and pay it back within a week. In our experiment, all subjects who bought the good had to pay for the good at the end of the experimental session.

Champ, Patricia A., Richard C. Bishop, Thomas C. Brown, and Daniel W. McCollum. 1997. Using donation mechanisms to value nonuse benefits from public goods. *Journal of Environmental Economics and Management* 33:151–62.

Cummings, Ronald G., Steven Elliott, Glenn W. Harrison, and James Murphy. 1997. Are hypothetical referenda incentive compatible? *Journal of Political Economy* 105:609–21.

Cummings, Ronald G., Glenn W. Harrison, and Elisabet E. Rutström. 1995. Homegrown values and hypothetical surveys: Is the dichotomous choice approach incentive-compatible? *American Economic Review* 85:260–6.

Diamond, Peter A., and Jerry A. Hausman. 1994. Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives* 8:45–64.

Greene, William H. 1993. *Econometric analysis.* 2nd edition. New York: Macmillan.

Hanemann, W. Michael. 1994. Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8:19–43.

Johannesson, Magnus, Bengt Liljas, and Per-Olov Johansson. 1998. An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions. *Applied Economics.* In press.

Loomis, John, Thomas Brown, Beatrice Lucero, and George Peterson. 1996. Improving validity experiments of contingent valuation methods: Results of efforts to reduce the disparity of hypothetical and actual willingness to pay. *Land Economics* 72:450–61.

Mitchell, Robert C., and Richard T. Carson. 1989. *Using surveys to value public goods: The contingent valuation method.* Baltimore: Johns Hopkins University Press.

Nape, Steven, Peter Frykblom, Glenn W. Harrison, and James C. Lesley. 1995. Hypothetical bias and willingness to accept. Economics Working Paper No. B-95-09, College of Business Administration, University of South Carolina.

Neill, Helen R., Ronald G. Cummings, Philip T. Ganderton, Glenn W. Harrison, and Thomas McGuckin. 1994. Hypothetical surveys, provision rules, and real economic commitments. *Land Economics* 70:145–54.

Newbold, Paul. 1991. *Statistics for business and economics.* 3rd edition. Englewood Cliffs, NJ: Prentice-Hall.

Southern Economic Journal 1998, 65(1), 178–181

Notes and Comments

Imperfect Labor Mobility and Unemployment in LDCs: Comment

John Gilbert* and Mia Mikic†

1. Introduction

In a recent paper in this journal, Parai and Beladi (1997; henceforth PB) extend the well-known Harris-Todaro model to consider imperfect labor mobility. They find unambiguous support for protection in the case of a small open economy with unemployment and imperfect labor mobility (Proposition VII). We show that the PB conclusion that the optimal tariff is always positive is incorrect. Furthermore, this note contrasts PB's model with the existing literature on the open economy with unemployment. We show that in the limiting case of perfect labor mobility, their model does not produce results that economic intuition and the existing literature would lead us to expect. We believe that their specification of the elasticity of labor mobility has no economic rationale within the context of the Harris-Todaro framework. We briefly explore a simple alternative specification.

2. The Parai-Beladi Analysis

PB introduce a two-sector (urban/rural), two-factor (labor and capital), dual, constant-returns-to-scale economy in which capital is perfectly mobile within the domestic economy. Like Corden and Findlay (1975) and Batra and Naqvi (1987), PB also incorporate labor unemployment specific to the urban sector by means of the Harris-Todaro expected wage mechanism. Their contribution is to introduce the issue of labor mobility by utilizing a simplified version of Casas's (1984) specification of the elasticity of labor mobility. The key result of PB (Proposition VII) is that imperfect labor mobility in the presence of unemployment always implies a positive optimal tariff on the output of the urban sector (the importable) for a small economy.

In the case of a small open economy with perfect labor and capital mobility and sector-specific unemployment of the Harris-Todaro category, it has been shown that a small import tariff on the output of the urban sector causes a welfare loss because, in addition to the usual production and consumption distortions, it also increases the rate of unemployment in the urban sector. This result was first geometrically explored in Corden and Findlay (1975) and has been rigorously proven by Batra and Naqvi (1987).

---

* Department of Economics, The University of Auckland, Private Bag 92019, Auckland, New Zealand; E-mail j.gilbert@auckland.ac.nz; corresponding author.

† Department of Economics, The University of Auckland, Private Bag 92019, Auckland, New Zealand; E-mail m.mikic@auckland.ac.nz.

We are indebted to Ed Tower and an anonymous referee for helpful comments.

PB state that they "extend the Casas model further by incorporating unemployment caused by sector-specific sticky wages of the Harris-Todaro variety." This means in the limiting case of perfect labor mobility ($\epsilon = \infty$, where ϵ is the elasticity of labor mobility), their model must converge to the open economy Harris-Todaro model with mobile capital. It is clear that PB believe this to be the case. Therefore, we should be able to utilize the results obtained by Batra and Naqvi (1987) to verify the validity of their specification. According to the specification of PB (Equation 9), the rural/urban wage ratio is related to the ratio of employment in the two sectors in the following manner:

$$(W_1/\bar{W}_2) = (L_1/L_2)^{1/\epsilon}, \tag{1}$$

where W_1 is the wage in rural sector, $\bar{W}_2$ is the minimum wage in urban sector, and L_1 and L_2 denote employment in the rural and urban sectors, respectively. This specification implies that rural/urban migration takes place in response to changes in nominal wages, with the parameter ϵ determining the degree of mobility. However, as Harris and Todaro (1970) state, the distinguishing feature of the Harris-Todaro model is that "migration proceeds in response to urban/rural differences in *expected earnings,* with the urban employment rate acting as an equilibrating force on such migration" (p. 126, original emphasis). To illustrate the inconsistency of PB's specification, when ϵ in Equation 1 approaches infinity, equalization of rural/urban wage rates must occur (assuming that employment in both sectors is always positive):

$$\lim_{\epsilon \to \infty} W_1 = \bar{W}_2.$$

This implies unemployment must disappear altogether. To see why this is the case, note PB's Equation 10:

$$(L_1/L_2)^{1/\epsilon} = [L_2/(L_2 + L_u)] \tag{2}$$

where L_u is unemployment in the urban sector. Equation 2 implies that

$$\lim_{\epsilon \to \infty} L_u = 0.$$

Hence, the unemployment that PB state always rises with a tariff (Proposition V) cannot even exist when labor is perfectly mobile since the rate of urban unemployment is fixed at zero. This result is in stark contrast to the results obtained in the cited literature, where perfect labor mobility cannot work to eliminate urban unemployment, and we are unable to see the economic rationale underlying PB's result. As we illustrate in the following section, the probability of employment in the urban sector plays a key role in models of the Harris-Todaro variety, and it is not reasonable to have it fixed, as the foregoing specification implies.[1] We propose the following specification of a model to incorporate the concept of imperfect labor mobility in a way that is consistent with the standard Harris-Todaro theoretical framework.

---

[1] The unconventional outcomes of PB continue to hold even without PB's simplification of Casas's (1984) specification of labor immobility. Casas's (1984, p. 750) Equation 5 has

$$(L_1/L_2) = k[1/b(\bar{W}_2/W_1)]^{\epsilon}.$$

PB assume that $k = b = 1$, from which they obtain their Equations 9 and 10. However, even had they set $b < 1$, their results would still not converge to the standard results. Unemployment would exist, but the specification would still imply that the urban unemployment rate is fixed when labor is perfectly mobile.

3. An Alternative Specification

We assume a small, dual, open economy where the rural sector is labeled 1 and the urban sector labeled 2. We make the standard assumptions that labor and capital are in fixed supply, markets are perfectly competitive, and returns to scale are constant. We define the cost of migration, ρ, as[2]

$$\rho = \bar{W}_2\pi - W_1, \tag{3}$$

where $\pi \equiv L_2/L_C = L_2/(L_2 + L_U)$ is the probability of obtaining employment in the city.[3] Migration from the rural to the urban sector stops when, for the marginal worker, the rural wage plus the cost of migration for that worker is equalized to the expected urban wage.[4] We now introduce what we term the elasticity of labor migration, ϵ, which we define as

$$\hat{L}_C = \epsilon\hat{\rho},$$

where a "^" represents a proportional change and $0 \leq \epsilon \leq \infty$. This elasticity is the change in the proportion of labor choosing to live in the city induced per proportional change in the expected wage differential. In the limiting case where $\epsilon = \infty$ (or where ρ is fixed for each worker), this specification replicates the Harris-Todaro results. Since from the proportional change form of the definition of π we have $\hat{\pi} = \hat{L}_2 - \hat{L}_C$, ϵ can be related to $\hat{\pi}$ in the following manner:

$$\hat{\pi} = \hat{L}_2 - \epsilon\hat{\rho}.$$

In the neighborhood of free trade, the change in utility as one worker moves from the rural to the urban sector is given by

$$dU = (W_1 + \rho)dL_1 + \bar{W}_2 dL_2 + RdK_1 + RdK_2,$$

where R is the rental rate on mobile capital and the marginal utility of income is defined to be unity. Since

$$dL_2 = L_C d\pi + \pi dL_C,$$

and making use of the fact that $dK_1 + dK_2 = 0$, the expression for the change in utility can be expressed after some manipulation as

$$dU = \bar{W}_2 L_C d\pi. \tag{4}$$

This equation forms the foundation of our results, and it is intuitively obvious. When one worker migrates at initial prices holding π constant, the value of output at world prices net of migration cost is unchanged and therefore so is utility. So the change in utility must be given by the value at initial prices of the extra output produced by the initial urban labor force when π rises, and that is given by Equation 4.

---

[2] The cost of migration is a variable that may depend on, for example, each worker's locational preferences, attachments to existing arrangements, or the high cost of relocation, as per PB (p. 180).

[3] To simplify the notation, we let the symbol L_c (labor in the city) represent the total of urban employed and unemployed.

[4] We assume no "commuting."

4. Conclusion

Equation 4 illustrates that even with imperfect labor mobility, the sign of the change in the probability of employment in the urban sector determines the sign of the change in utility. To determine the sign of the welfare effect of a small tariff, it is only necessary to determine what happens to the probability of urban employment (since the deadweight loss is of second-order magnitude). It can be shown that when labor is completely immobile, while capital is perfectly mobile, a small tariff on imports of the urban good will increase the probability of employment in the urban sector (thus $d\pi$ is positive). A small tariff therefore raises welfare for a small economy with unemployment, as PB indicate. However, unemployment does not rise. The tariff draws labor into urban employment solely from the urban pool of unemployed (in contrast to PB's Proposition V). Of course, a tariff is not a first-best instrument in any case since it imposes a by-product distortion by limiting imports.

When labor is perfectly mobile, a small tariff on urban output will have an adverse effect on the likelihood of employment in that sector ($d\pi$ is negative), and thus welfare falls with the imposition of an import tariff. This is the conventional result proven by Batra and Naqvi (1987). The tariff with the urban wage frozen raises the urban rental rate on capital. With capital perfectly mobile, this raises the rural rental by the same proportion. With the price of the rural output fixed by the perfectly elastic world excess demand, this shrinks the rural wage. With a given ρ and the urban wage fixed, from Equation 3 π must decline. In general, however, $d\pi$ is of ambiguous sign and does not allow for any definite conclusions about the welfare effects of a small tariff in a small economy with unemployment and imperfect labor mobility.[5] Therefore, we believe that PB overstate their case for protectionism. Only when the elasticity of labor migration is sufficiently close to zero is a tariff (or output subsidy) welfare enhancing.

References

Batra, R. N., and N. Naqvi. 1987. Urban unemployment and gains from trade. *Economica* 54:381–95.

Casas, F. R. 1984. Imperfect factor mobility: A generalization and synthesis of two-sector models of international trade. *Canadian Journal of Economics* 17(4):747–61.

Chao, Chi-Chur, and E. S. H. Yu. 1994. Urban growth, externality and welfare. *Regional Science and Urban Economics* 24:565–76.

Corden, W. M., and R. Findlay. 1975 Urban unemployment, intersectoral capital mobility and development policy. *Economica* 42:59–78.

Harris, J. R., and M. P. Todaro. 1970. Migration, unemployment and development: A two-sector analysis. *American Economic Review* 60:126–42.

Parai, A. K., and H. Beladi. 1997. Imperfect labor mobility and unemployment in LDCs. *Southern Economic Journal* 64:180–90.

[5] The solution for the proportional change in the probability of urban employment in this model is

$$\hat{\pi} = \frac{-\lambda_{L1}(\sigma_1\lambda_{K1}\lambda_{LC} + \sigma_2\theta_{L1}\lambda_{K1}\lambda_{LU}) + \epsilon W_1\theta_{K1}\lambda_{LC}(\lambda_{L1} - \lambda_{K1})}{-(\theta_{L1}\theta_{K1}\lambda_{K1}\lambda_{L1}\lambda_{LU}) + \epsilon\bar{W}_2\pi\lambda_{LC}(\lambda_{K1} - \lambda_{L1})}\cdot\hat{p}$$

where λ_{ij} is the proportion of factor i employed in sector j, θ_{ij} is the cost share of factor i in sector j, σ_j is the elasticity of substitution between capital and labour in sector j, and p is the relative price of the urban output. Note that by the stability condition under perfect labor mobility $(\lambda_{L1} - \lambda_{K1}) > 0$ (see Chao and Yu [1994]), the second term in the numerator is positive and the second term in the denominator negative, leading to the ambiguity in results.

Southern Economic Journal 1998, 65(1), 182–184

Imperfect Labor Mobility and Unemployment in LDCs: Reply

Amar K. Parai* and Hamid Beladi†

In a recent paper (Parai and Beladi 1997; PB hereafter), we have analyzed the implications of growth and trade policies for a small open economy facing imperfect labor mobility and unemployment of the Harris-Todaro variety. We have used the Casas (1984) specification of the labor *immobility* phenomenon for a Harris-Todaro type economy, and have shown that most of the results in Harris-Todaro framework remain unaltered even under imperfect labor mobility, provided that the elasticity of labor mobility parameter exceeds a critical minimum value. On the optimal tariff issue, Gilbert and Mikic (1997; GM hereafter) find our results counterconventional. In GM's view, the nonconventional result in PB is due to our simplification of the labor mobility specification given by Casas. In this note, we offer our response to GM's comments.

First of all, we appreciate GM's attention to our apparently counterconventional result about the positive optimal tariff because their comment has given us an opportunity to rectify a minor algebraic error in one of our results and clear up the confusions in its policy implications. But their conclusion that it is because of our simplification assumption of unitary values of the scale parameters a and b in Casas's specification is incorrect. Had we maintained the Casas specification intact, PB's equation 10 would become

$$[b/(a)^{1/\epsilon}](L_1/L_2)^{1/\epsilon} = [L_2/(L_2 + L_u)]. \tag{1}$$

Logarithmic differentiation of Equation 1 would again yield equations 28 and 37 in PB. Consequently, the nonunitary values of a and b would leave all the comparative static results in PB unchanged.

On closer scrutiny of our algebra, however, we find an inadvertent error in the derivation of the optimal tariff, $\hat{t}$. The corrected expression reads

$$\hat{t} = a_s[1 - \psi) - G]/[a_s(\psi + G) + a_d],$$

where $\psi > 1$, $a_s \equiv (dX_2/dp)(p/E_2) > 0$, $a_d \equiv -(dD_2/dp)(p/E_2) > 0$, and $G \equiv [1/\{1 - \lambda_{L1}(1 + \epsilon)\}] < 0$. Thus, the sign of $\hat{t}$ becomes indeterminate as anticipated by GM.

As in Batra and Naqvi (1987), in PB as well, when labor mobility is perfect, a reduction of tariff will increase the social welfare of a Harris-Todaro economy. This can be verified from PB's equation 41. As $\epsilon \to \infty$, the third term within the bracket in the right-hand side of equation 41 disappears, and $(dU/dt) < 0$. In a Harris-Todaro economy, rural-to-urban migration of labor causes a distortion in the output structure. Because of the migration, agricultural output is less than what would have been otherwise. Specifically, at the equilibrium production point, the

* Department of Economics, SUNY College at Fredonia, Fredonia, NY 14063; corresponding author.
† Department of Economics and Finance, University of Dayton, Dayton, OH 45469-2240, USA; E-mail beladi@udayton.edu.

social marginal cost of producing the industrial good becomes higher than its private marginal cost (for details, see Parai and Batra 1987, p. 316). Naturally, any policy that increases the output of agricultural sector would reduce the adverse effect of the production distortion and thereby increase welfare. A reduction of tariff on manufactured goods produced in the urban sector would be such a policy, and that is precisely the rationale behind the conventional argument against protection in the Harris-Todaro economy. In the literature, this point has been emphasized over and over again (see, e.g., Corden and Findlay 1975; Parai and Batra 1987; Batra and Naqvi 1987). Now why does this conventional result appear to be a paradox?

In the factor market distortions literature, the output structure is biased against the sector that pays the exogenously given higher wages to labor it employs (see, e.g., Batra 1973, chap. 10). Even in models where the wage differential is endogenous because of, say, imperfect mobility of labor, the same type of production bias is observed against the output of the sector that pays the wage differential (see, e.g., Casas 1984; Yu and Parai 1989). Consequently, any policy that increases the output of this sector will reduce the distortionary output effect and move the output structure closer to what would have been under nondistortionary situation. A tariff protection of, or a production subsidy to, the sector paying the wage differential thus becomes the second-best optimal policy. In the Harris-Todaro economy, there is an exogenous wage differential paid by the urban manufacturing sector. Then, following the policy recommendations of the factor market distortions literature, the manufacturing sector should be protected by a second-best tariff or production subsidy.

Thus, in the standard Harris-Todaro model, two apparently different distortions call for diametrically opposite policy recommendations. The antiprotectionist trade policy for a standard Harris-Todaro economy apparently signifies the dominance of the migration-related distortion over the wage differential–based distortion, and that shows the importance of the unique migration phenomenon, which happens to be the novelty of the Harris-Todaro model.

In PB's model, we distort the conventional Harris-Todaro economy even further by introducing imperfect labor mobility. The wage rate in the manufacturing sector is institutionally set at a higher rate than its counterpart in the rural sector. But workers in the agricultural sector do have a preference of living in rural areas over migrating to urban centers of glitter and glee. As a result, migration of labor toward the higher-wage sectors is limited. Because of the limited migration, the output of manufacturing (agricultural) sector would not increase (decrease) by as much as it would under perfect labor mobility. Thus, imperfect labor mobility dampens the strength of the migration-related distortion in Harris-Todaro and causes some indeterminate results.

GM raise questions about our use of the Casas (1984) specification in the context of Harris-Todaro unemployment. Specifically, in the presence of unemployment as labor mobility becomes perfect, the Casas specification yields

$$\lim_{\epsilon \to \infty} (W_1/\bar{W}_2) = b = \lim[L_2/(L_2 + L_u)].$$

This implies that with $b < 1$, industrial employment (L_2) becomes a constant fraction of the urban labor force ($L_2 + L_u$) at the equilibrium when migration stops. This is no less "reasonable" an implication than a positive constant differential (ρ) between the expected urban and rural wage rates at the equilibrium in GM's alternative specification (eq. 5 in GM) of the Harris-Todaro model.

References

Batra, R. N. 1973. *Studies in pure theory of international trade.* New York: St. Martin's Press.

Batra, R. N., and N. Naqvi. 1987. Urban unemployment and gains from trade. *Economica* 54:381–95.

Casas, F. R. 1984. Imperfect factor mobility: A generalization and synthesis of two-sector models of international trade. *Canadian Journal of Economics* 17:747–61.

Corden, W. M., and R. Findlay. 1975. Urban unemployment, intersectoral capital mobility and development policy. *Economica* 42:59–78.

Gilbert, J., and M. Mikic. 1997. Imperfect labor mobility and unemployment in LDCs: Comment. Mimeograph.

Harris, J. R., and M. P. Todaro. 1970. Migration, unemployment and development: a two-sector analysis. *American Economic Review* 60:126–42.

Parai, A. K., and R. N. Batra. 1987. Customs union and unemployment in LDC's. *Journal of Development Economies* 26:311–9.

Parai, A. K., and H. Beladi. 1997. Imperfect labor mobility and unemployment in LDC's. *Southern Economic Journal* 64:180–90.

Yu, E. S. H., and A. K. Parai. 1989. Factor immobility and gains from trade. *Southern Economic Journal* 55:601–9.