Title
Three essays on behavioral economics

Author
Meng, Juanjuan

Publication Date
2010

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Three Essays on Behavioral Economics

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Economics

by

Juanjuan Meng

Committee in charge:

Professor Vincent P. Crawford, Chair
Professor Julie B. Cullen
Professor Uri Gneezy
Professor Joel Sobel
Professor Allan Timmermann

2010

The Dissertation of Juanjuan Meng is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2010

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Vincent P. Crawford for his support as the chair of my committee. Through multiple drafts and many long nights, his insightful guidance and general support have proved to be invaluable.

Chapter 1, in full, is currently being prepared for publication as a forthcoming paper in *American Economic Review.* Crawford, Vincent P.; Meng, Juanjuan. The dissertation author was one of the primary authors of this paper.

Chapter 2, in full, has been submitted for publication of the material. Meng, Juanjuan. The dissertation author was the primary author of this material.

VITA

2005	Bachelor of Arts, Peking University
2005-2006	Teaching Assistant, Department of Economics
	University of California, San Diego
2006	Master of Arts, University of California, San Diego
2006-2010	Teaching Assistant University of California, San Diego
2010	Doctor of Philosophy, University of California, San Diego

PUBCLICATIONS

"New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Utility with Targets for Hours and Income", co-authored with Vincent P. Crawford (Forthcoming *American Economics Review*)

FIELDS OF STUDY

Major Field: Economics

Studies in Behavioral Economics
Professors Vincent P. Crawford, Uri Gneezy and Joel Sobel

Studies in Applied Microeconomics
Professor Julie B. Cullen

Studies in Behavioral Finance
Professor Allan Timmermann

ABSTRACT OF THE DISSERTATION

Three Essays on Behavioral Economics

by

Juanjuan Meng

Doctor of Philosophy in Economics

University of California, San Diego, 2010

Professor Vincent P. Crawford, Chair

This dissertation investigates two topics on behavioral economics: reference-dependent preferences and social utility. Chapter 1 and 2 provide field evidence from labor market and financial market to support reference-dependent model that treats

expectations as reference points. Chapter 3 explores the implications of social distance on the endogenous emergence of personal relationships and impersonal market exchange.

Chapter 1: A model of cabdrivers' labor supply is proposed, building on Farber's (2005, 2008) empirical analyses and Kőszegi and Rabin's (2006; henceforth "KR") theory of reference-dependent preferences. Following KR, the proposed model has targets for hours as well as income, determined by proxied rational expectations. The model, estimated with Farber's data, reconciles his finding that stopping probabilities are significantly related to hours but not income with Colin Camerer et al.'s (1997) negative "wage" elasticity of hours; and avoids Farber's criticism that estimates of drivers' income targets are too unstable to yield a useful model of labor supply.

Chapter 2: An investor' aversion to losses relative to a reference point in the stock market predicts a V-shaped relationship between the optimal position in a stock and current gains from that stock. Estimates from Odean's (1999) individual trading records show that (i) the predicted V-shape relationship exists for a large majority of investors, and (ii) expectations are the most likely determinant of investors' reference points. The V-shaped relationship and the implication of the initial purchase decision that expectations are mostly positive yield a simple explanation of the disposition effect.

Chapter 3: Personal relationships and impersonal exchange have been previously modeled in ways that prevent them from coexisting in equilibrium as contract enforcement mechanisms. Empirical evidence nonetheless suggests that they

sometimes coexist. This paper introduces social surplus into exchange payoff, which is determined by social distance and specific to personal relationships but not to impersonal exchange. This approach allows the two modes of exchange to coexist in equilibrium. The possibility of impersonal exchange improves welfare and equality among buyers in general. But there also exist cases where competition between the two forms of exchange makes welfare and equality deteriorate.

Chapter 1

New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income

1.1. Introduction

In the absence of large income effects, a neoclassical model of labor supply predicts a positive wage elasticity of hours. However, Camerer et al. (1997) collected data on the daily labor supply of New York City cabdrivers, who unlike most workers in modern economies are free to choose their own hours, and found a strongly negative elasticity of hours with respect to their closest analog of a wage, realized earnings per hour. In Camerer et al.'s dataset, realized earnings per hour (which they call the "wage") is uncorrelated across days but positively serially correlated within a day, so that high earnings early in a day signal higher earnings later that day, and a neoclassical model predicts a positive elasticity even though realized earnings per hour is not precisely a wage. If instead realized earnings per hour is serially uncorrelated within a day, as Farber (2005) shows is roughly true in his dataset (see however our analysis in Section 1.3.3), then a driver with high early earnings experiences a small change in income but no change in expected wage, and a neoclassical model predicts an elasticity near zero. A neoclassical model could only explain Camerer et al.'s strongly negative elasticities via an implausibly large negative serial correlation of realized earnings per hour.

To explain their results, Camerer et al. informally proposed a model in which drivers have daily income targets and work until the target is reached, and so work less on days when earnings per hour are high. Their explanation is in the spirit of Daniel Kahneman and Amos Tversky's (1979) and Tversky and Kahneman's (1991) prospect theory, in which a person's preferences respond not only to income as usually assumed, but also to a reference point; and there is "loss aversion" in that the person is more sensitive to changes in income below the reference point ("losses") than changes above it ("gains"). If a driver's reference point is a daily income target, then loss aversion creates a kink that tends to make realized daily income bunch around the target, so that hours have a negative elasticity with respect to realized earnings per hour.

Farber (2008, p. 1069) suggests that a finding that labor supply is reference-dependent would have significant policy implications:

> Evaluation of much government policy regarding tax and transfer programs depends on having reliable estimates of the sensitivity of labor supply to wage rates and income levels. To the extent that individuals' levels of labor supply are the result of optimization with reference-dependent preferences, the usual estimates of wage and income elasticities are likely to be misleading.

Although Camerer et al.'s analysis has inspired a number of empirical studies of labor supply, the literature has not yet converged on the extent to which the evidence supports reference-dependence.[1] Much also depends on reference-

[1] KR (2006) and Farber (2008) survey some of the empirical literature. Gerald S. Oettinger's (1999)

dependence's scope and structure: If it were limited to inexperienced workers or unanticipated changes, its direct relevance to most policy questions would be small, though it might still have indirect policy relevance via its influence on the structure of labor relationships. This paper seeks to shed additional light on these issues, building on two recent developments: Farber's (2005, 2008) empirical analyses of cabdrivers' labor supply and KR's (2006; see also 2007, 2009) theory of reference-dependent preferences.

Farber (2005) collected and analyzed data on the labor supply decisions of a new set of New York City cabdrivers. He found that, before controlling for driver fixed effects, the probability of stopping work on a day is significantly related to realized income that day, but that including driver fixed effects and other relevant controls renders this effect statistically insignificant.

Farber (2008) took his 2005 analysis a step further, introducing a structural model based on daily income targeting that goes beyond the informal explanations in previous empirical work. He then estimated a reduced form, treating drivers' income targets as latent variables with driver-specific means and driver-independent variance, both assumed constant across days of the week—thus allowing the target to vary across days for a given driver, but only as a random effect. Constancy across days of the week is violated in the sample, where Thursdays' through Sundays' incomes are systematically higher than those of other days, and the hypothesis that income is

---

field study found increased daily participation by stadium vendors on days on which the anticipated wage was higher, as suggested by the neoclassical model, and in seeming contrast to Camerer et al.'s finding of a negative response of hours to (partly unanticipated) increases in wage. Ernst Fehr and Lorenz Goette's (2007) field experiment found increased participation by bicycle messengers, but reduced effort, in response to announced increases in their commission. They argued that effort is a more accurate measure of labor supply and concluded that the supply of effort is reference-dependent.

constant across days of the week is strongly rejected (*p*-value 0.0014, F-test with robust standard errors). Farber included day-of-the-week dummies in his main specifications for the stopping probability, but this turns out to be an imperfect substitute for allowing the mean income target to vary across days of the week.

Farber found that a sufficiently rich parameterization of his targeting model fits better than a neoclassical model, and that the probability of stopping increases significantly and substantially when the target is reached; but that his model cannot reconcile the increase in stopping probability at the target with the smooth aggregate relationship between stopping probability and realized income. Further, the estimated random effect in the target is large and significantly different from zero, but with a large standard error, which led him to conclude that the targets are too unstable to yield a useful reference-dependent model of labor supply (p. 1078):

> There is substantial inter-shift variation, however, around the mean reference income level.… To the extent that this represents daily variation in the reference income level for a particular driver, the predictive power of the reference income level for daily labor supply would be quite limited.

KR's (2006) theory of reference-dependent preferences is more general than Farber's (2008) model in most respects, but takes a more specific position on how targets are determined. In KR's theory applied to cabdrivers, a driver's preferences reflect both the standard consumption utility of income and leisure and reference-dependent "gain-loss" utility, with their relative importance tuned by a parameter. As in Farber's model, a driver is loss-averse; but he has a daily target for hours as well as income, and working longer than the hours target is a loss, just as earning less than the income target is. Finally, KR endogenize the targets by setting a driver's targets equal to his theoretical rational expectations of hours and income, reflecting the belief that drivers in steady state have learned to predict their distributions.[2]

This paper uses Farber's (2005, 2008) data to reconsider the reference-dependence of cabdrivers' labor supply, adapting his econometric strategies to estimate models based on KR's (2006) theory. Section 1.2 introduces the model. Following KR, we allow for consumption as well as gain-loss utility and hours as well as income targets; but when we implement the model we follow Farber (2008) in assuming that drivers are risk-neutral in income.

To complete the specification, we must describe how a driver's targets are determined and, for some of our analysis, how he forms his expectations about earnings hour by hour during a day. In an important departure from Farber's approach, we follow KR in conceptualizing drivers' targets and expected earnings as rational expectations, but for simplicity we depart from KR in treating them as point expectations rather than distributions. Because KR's model has no errors, their distributions are necessary for the existence of deviations from expectations, without which their model reduces to a neoclassical model. Our model has errors and so has deviations even with point expectations. We operationalize the targets and expected earnings via natural sample proxies with limited endogeneity problems as explained below, for expected earnings assuming for simplicity that earnings per hour are

[2] In theory there can be multiple expectations that are consistent with the individual's optimal behavior, given the expectations. KR use a refinement, "preferred personal equilibrium," to focus on the self-confirming expectations that are best for the individual. Most previous analyses have identified targets with the status quo; but as KR note, most of the available evidence does not distinguish the status quo from expectations, which are usually close to the status quo. Even so, our analysis shows that KR's rational-expectations view of the targets has substantive implications for modeling cabdrivers' labor supply. KR's view of the targets has also been tested and supported in laboratory experiments by Johannes Abeler et al. (2011).

serially uncorrelated within a day (as well as across days). This last simplification is motivated by Farber's (2005) finding, in a detailed econometric analysis of his dataset, of only a weak and insignificant relationship, which led him to argue that hourly earnings are so variable and unpredictable that "predicting hours of work with a model that assumes a fixed hourly wage rate during the day does not seem appropriate." We note however that Camerer et al. did find some within-day predictability of earnings in their dataset. It also seems plausible that drivers on the ground may be able to predict their earnings better than even the most careful econometrics. Given no serial correlation and risk-neutrality in income, and ignoring option value in the stopping decision, a driver's expected hourly earnings are equivalent to a predetermined (though random) daily schedule of time-varying wages.[3]

If the weight of gain-loss utility is small, our model mimics a neoclassical labor-supply model, so that the elasticity of hours with respect to earnings per hour is normally positive. If the weight of gain-loss utility is large, perfectly anticipated changes in earnings per hour still have this neoclassical implication because gain-loss utility then drops out of a driver's preferences. However, *un*anticipated changes may then have non-neoclassical implications. In particular, when the income target has an important influence on a driver's stopping decision, a driver who values income but is "rational" in the reference-dependent sense of prospect theory will tend to have a

[3] Farber (2008) modeled a driver's stopping decision by estimating a daily latent income target and continuation value, assuming that a driver stops working when his continuation value falls below the cost of additional effort. He defined continuation value to include option value; but if option value is truly important, his linear specification of continuation value is unlikely to be appropriate. We simply assume that drivers' decisions ignore option value, as Thierry Post et al. (2008) did, and as seems behaviorally reasonable. Farber's (2008) and our treatments of drivers' decisions are both first-order proxies for globally optimal stopping conditions that depend on unobservables, which treatments both yield coherent results, despite their flaws.

negative elasticity of hours with respect to earnings per hour, just as Camerer et al. found.

Section 1.3 reports our econometric estimates. In Section 1.3.1 we estimate probit models of the probability of stopping with an index function that is linear in cumulative shift hours and income as in Farber (2005), but splitting the sample according to whether a driver's earnings early in the day are higher or lower than his proxied expectations. This "early earnings" criterion should be approximately uncorrelated with errors in the stopping decision, limiting sample-selection bias. To avoid confounding due to our operationalization of the targets being partly determined by the variables they are used to explain, we proxy drivers' rational point expectations of a day's income and hours, driver/day-of-the-week by driver/day-of-the-week, by their sample averages up to but not including the day in question.

In a neoclassical model, when earnings per hour is serially uncorrelated within a day, it is approximately irrelevant whether early earnings are unexpectedly high or low, because this affects a driver's income but not his expected earnings later in the day, and the income effect is negligible. But in a reference-dependent model, high early earnings make a driver more likely to reach his income target before his hours target, and this has important consequences for behavior. In our estimates drivers' stopping probabilities happen to be more strongly influenced by the second target a driver reaches on a given day than by the first. As a result, when early earnings are high, hours (but not income) has a strong and significant effect on the stopping probability, either because the driver reaches his hours target or because his marginal utility of leisure increases enough to make additional work undesirable. When early earnings are low, this pattern is reversed.[4] Such a reversal is inconsistent with a neoclassical model, in which the targets are irrelevant; but it is gracefully explained by a reference-dependent model. If preferences were homogeneous, as Farber's and our models assume, drivers' stopping probabilities would either all tend to be more strongly influenced by the first target reached on a given day, or all by the second. Thus the pattern of significance in our results is one of the two that are characteristic of a reference-dependent model with homogeneous preferences, and as such is powerful evidence for reference-dependence, even though with heterogeneous preferences other patterns are possible.[5]

Further, because the elasticity of hours with respect to earnings per hour is substantially negative when the income target is the dominant influence on stopping probability, but near zero when the hours target is dominant, and on a typical day some drivers' earnings are higher than expected and others' lower, KR's distinction between anticipated and unanticipated wage changes can easily reconcile the presumably normally positive incentive to work of an anticipated increase in earnings per hour, with a negative observed aggregate elasticity of hours. As KR put it (2006, p. 1136):

> In line with the empirical results of the target-income literature, our model predicts that when drivers experience unexpectedly high wages in the morning, for any given afternoon wage they are less likely to continue work. Yet expected wage increases will tend to increase both willingness to show up to work, and to drive in the afternoon once there. Our model therefore replicates the key insight of the literature that

---

[4] Our estimates reverse the patterns of significance from the analogous results in Table 2 of the original version of this paper, Crawford and Meng (2008), suggesting that those results were biased due to the endogeneity of the sample-splitting criterion we used there: whether realized earnings were higher or lower than the full-sample average for a given driver and day-of-the-week.

[5] Kirk Doran (2009), in an important study of yet another group of New York City cabdrivers, with enough data to estimate individual-level effects, finds considerable heterogeneity in drivers' behavior, with some reference-dependent and others neoclassical.

> exceeding a target income might reduce effort. But in addition, it both provides a theory of what these income targets will be, and—through the fundamental distinction between unexpected and expected wages—avoids the unrealistic prediction that generically higher wages will lower effort.

Finally, with a distribution of earnings the model can also reproduce Farber's (2005) findings that aggregate stopping probabilities are significantly related to hours but not earnings, but nonetheless respond smoothly to earnings.

In Section 1.3.2 we use the pooled sample to estimate a reduced-form model of the stopping probability, with dummy variables to measure the increments due to hitting the income and hours targets as in Farber's (2008) Table 2 but with our proxied targets instead of Farber's estimated targets. The estimated increments are large and significant, again with a sign pattern strongly suggestive of a reference-dependent model, and the effects of income and hours come mostly from whether they are above or below their targets rather than from levels per se.

In Section 1.3.3 we use the pooled sample to estimate a structural reference-dependent model in the spirit of Farber's (2008) model, again with the changes suggested by KR's theory. In our model the weight of gain-loss utility and the coefficient of loss aversion are not separately identified, but a simple function of them is identified, and its estimated value deviates strongly and significantly from the value implied by a neoclassical model. There is more than enough independent variation of hours and income and our proxies for drivers' targets to identify our model's other behavioral parameters, and to distinguish bunching of realized hours due to targeting from bunching that occurs for conventional neoclassical reasons. The parameter estimates are plausible and generally confirm the conclusions of Section 1.3.1-2's analyses. The estimated model again implies significant influences of income and hours targets on stopping probabilities, in a pattern that is gracefully explained by a reference-dependent model but inconsistent with a neoclassical model; and resolves the puzzles left open by Farber's analyses.

Our results suggest that reference-dependence is an important part of the labor-supply story in Farber's dataset, and that using KR's model to take it into account does yield a useful model of cabdrivers' labor supply. The key aspect of our analysis, which allows it to avoid Farber's criticism that drivers' estimated targets are too unstable to yield a useful model, is implementing KR's rational-expectations view of drivers' income and hours targets by finding natural sample proxies that limit endogeneity problems, rather than estimating the targets as latent variables.

Section 1.4 is the conclusion.

1.2. The Model

This section introduces our model of cabdrivers' labor supply decisions.

Treating each day separately as in all previous analyses, consider the preferences of a given driver on a given day.[6] Let I and H denote his income earned and hours worked that day, and let I^r and H^r denote his income and hours targets for the day. We write the driver's total utility, $V(I, H|I^r,H^r)$, as a weighted average of consumption utility $U_1(I) + U_2(H)$ and gain-loss utility $R(I, H|I^r,H^r)$, with weights 1 –

[6] A driver sometimes works different shifts (day or night) on different days but never more than one a day. Given that drivers seem to form daily targets, it is natural to treat the shift, or equivalently the driver-day combination, as the unit of analysis.

η and η (where $0 \le \eta \le 1$), as follows:[7]

(1.1) $$V(I,H \mid I^r, H^r) = (1-\eta)(U_1(I)+U_2(H)) + \eta R(I,H \mid I^r, H^r)\,,$$

where gain-loss utility

(1.2) $$R(I,H \mid I^r, H^r) = 1_{(I-I^r \le 0)}\lambda(U_1(I)-U_1(I^r)) + 1_{(I-I^r > 0)}(U_1(I)-U_1(I^r))$$
$$+1_{(H-H^r \ge 0)}\lambda(U_2(H)-U_2(H^r)) + 1_{(H-H^r < 0)}(U_2(H)-U_2(H^r)).$$

Because to our knowledge this is the first test of KR's theory, for simplicity and parsimony (1.1)-(1.2) incorporate some assumptions KR made provisionally: Consumption utility is additively separable across income and hours, with $U_I(\cdot)$ increasing in I, $U_2(\cdot)$ decreasing in H, and both concave. In keeping with the "narrow bracketing" assumption that drivers evaluate consumption and gain-loss utility day by day, $U_I(I)$ is best thought of as a reduced form, including the future value of income not spent today. This suggests that the marginal utility of income is approximately constant and, treating $U_I(\cdot)$ as a von Neumann-Morgenstern utility function, that consumption utility is approximately risk-neutral in daily income, a restriction Farber (2008) and we impose in our structural analyses. Gain-loss utility is also separable, with its components determined by the differences between realized and target consumption utilities. As in a leading case KR often focus on (their Assumption A3'), gain-loss utility is a linear function of those utility differences, thus ruling out prospect theory's "diminishing sensitivity." Finally, losses have a constant weight λ relative to gains, "the coefficient of loss aversion," assumed to be the same for income and hours. This leaves open the question of whether preferences are reference-dependent in both

[7] KR (2006, 2007) use a different parameterization, in which consumption utility has weight 1 and gain-loss utility has weight η. Our η is a simple transformation of theirs.

income and hours. Estimates that allow λ to differ for income and hours robustly show no significant difference (although the estimated λ for hours is always larger than that for income), so in all but Section 1.3.3's structural estimation we assume for simplicity that λ is the same for both.

We follow KR in conceptualizing the income and hours targets I^r and H^r as rational expectations, but for simplicity, we assume that they are point expectations. This exaggerates the effect of loss aversion, and if anything biases the comparison against a reference-dependent model. We operationalize expectations via sample proxies with limited endogeneity problems as explained in Section 1.3. We further assume that the driver is approximately risk-neutral in daily income, so that only its expected value matters to him.

Our model allows a simple characterization of a driver's optimal stopping decision with a target for hours as well as income, which parallels Farber's (2005, 2008) characterization of optimal stopping with income targeting. To simplify this discussion, assume for the moment that the driver has a daily wage in the sense of predetermined daily expected earnings per hour w^e that are constant over time. Further assume that $\lambda \geq 1$, reflecting the almost universal empirical finding that there is loss rather than gain aversion.

The optimal stopping decision then maximizes $V(I, H|I^r, H^r)$ as in (1.1) and (1.2), subject to the linear menu of expected income-hours combinations $I = w^e H$. When $U_1(\cdot)$ and $U_2(\cdot)$ are concave, $V(I, H|I^r, H^r)$ is concave in I and H for any targets I^r and H^r. Thus the driver's decision is characterized by a first-order condition, generalized to allow kinks at the reference points: He continues if expected earnings per hour exceeds the relevant marginal rate of substitution and stops otherwise.[8] Table 1 lists the marginal rates of substitution in the interiors of the four possible gain-loss regions, expressed as hours disutility costs per unit of income. Under our assumptions that gain-loss utility is additively separable and determined component by component by the difference between realized and target consumption utilities, when hours and income are both in the interior of the gains or loss domain, the marginal rate of substitution is the same as for consumption utilities alone and the stopping decision follows the neoclassical first-order condition. But when hours and income are in the interiors of opposite domains, the marginal rate of substitution equals the consumption-utility trade-off times a factor that reflects the weight of gain-loss utility and the coefficient of loss aversion, $(1 - \eta + \eta\lambda)$ or $1/(1 - \eta + \eta\lambda)$. On boundaries between regions, where $I = I^r$ and/or $H = H^r$, the marginal rates of substitution are replaced by generalized derivatives whose left- and right-hand values equal the interior values.

Table 1.1. Marginal Rates of Substitution with Reference-Dependent Preferences

	Hours gain ($H < H^r$)	Hours loss ($H > H^r$)
Income gain ($I > I^r$)	$-U_2'(H)/U_1'(I)$	$-[U_2'(H)/U_1'(I)][1-\eta+\eta\lambda]$
Income loss ($I < I^r$)	$-[U_2'(H)/U_1'(I)]/[1-\eta+\eta\lambda]$	$-U_2'(H)/U_1'(I)$

[8] If a driver's expected wage varies too much within shift or in response to experience, his optimization problem may become non-convex, in which case optimal stopping requires more foresight than we assume. Further, more general specifications that allow diminishing sensitivity do not imply that $V(I, H|I^r, H^r)$ is everywhere concave in I and H. Although such specifications probably still allow an analysis like ours, as do other expectations formation rules, we avoid these complications.

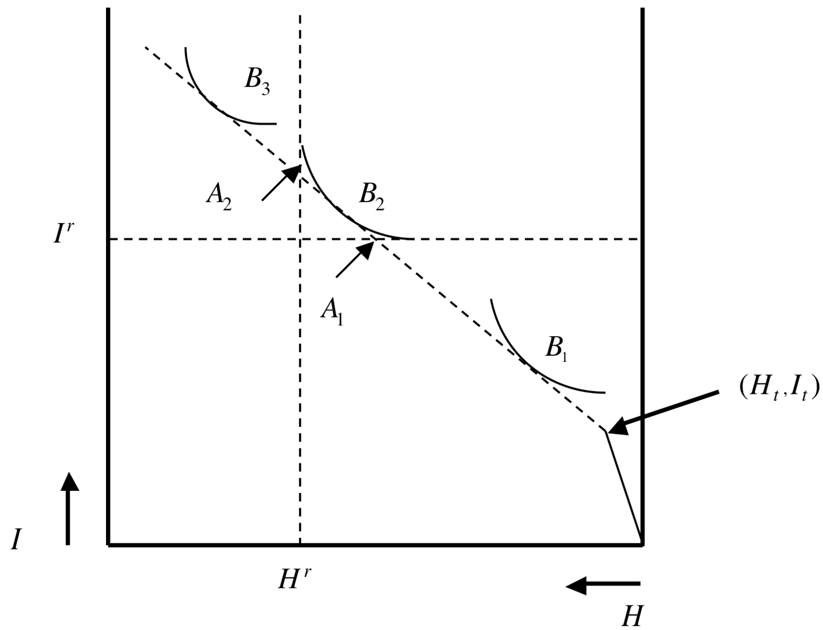


Figure 1.1. A Reference-dependent Driver's Stopping Decision when Realized Earnings are Higher than Expected

Figure 1.1, in which hours are measured negatively as a "bad," illustrates the driver's optimal stopping decision when, after an initial blip of higher than expected realized earnings, both realized and expected earnings per hour, w^a and w^e, remain constant and equal. As a result of the blip, total realized earnings remain higher than initially expected, and the income target is reached before the hours target. We stress that the constancy of w^e and w^a and the fact that there are no surprises after the blip are only for illustration; the important thing is that total realized earnings remain higher than expected. The case where realized earnings are lower than expected and the hours target is reached before the income target is completely analogous.

Letting I_t and H_t denote earnings and hours by the end of trip t, the driver starts in the lower right-hand corner with $(H_t, I_t) = (0, 0)$, followed by an initial period of higher than expected realized earnings. Total earnings and hours then increase along a weakly monotone path (not shown), heading northwest. The path is actually a step function, but because mean trip length is only 12 minutes (Farber (2005, Section V)), the path can be treated as smooth and I and H as continuous variables. After any given trip t, the driver anticipates moving along a line $I = w^eH$, starting from the current (H_t, I_t). As hours and income accumulate, a driver who continues working passes through a series of domains such that the hours disutility cost of income weakly increases, whichever target is reached first—a reflection of the concavity of $V(I, H|I^r,H^r)$ in I and H. The driver considers stopping after each trip, stopping (ignoring option value) when his current expected wage first falls below his current hours disutility cost of income. This myopia may lead the driver to deviate from KR's preferred personal equilibrium (footnote 3), although this can matter only in our structural estimation. The driver stops at a point that appears globally optimal to him, given his myopic expectations. This conclusion extends to drivers who form their expectations in more sophisticated ways, unless their expected earnings vary too much.

For example, in the income-loss/hours-gain ($I_t < I^r$, $H_t < H^r$) domain, the hours disutility cost of income is $-[U_2'(H_t)/U_1'(I_t)]/[1-\eta+\eta\lambda]$ from the lower left cell of Table 1.1. Because in this domain hours are cheap relative to income ($(1 - \eta + \eta\lambda) \geq 1$ when $0 \leq \eta \leq 1$ and $\lambda \geq 1$), the comparison with expected earnings per hour favors working more than the neoclassical comparison. The indifference curves in Figure 1.1 with tangency points B_1, B_2, and B_3 represent alternative possible income-hours trade-offs for consumption utility, ignoring gain-loss utility. If a driver stops in the income-loss/hours-gain domain, it will be (ignoring discreteness) at a point weakly between B_1 and A_1 in the figure, where B_1 maximizes consumption utility on indifference curve 1 subject to $I = w^e H$ and A_I represents the point where the income target is reached. (The closer η is to one and the larger is $\lambda \geq 1$, other things equal, the closer the stopping point is to A_I.)

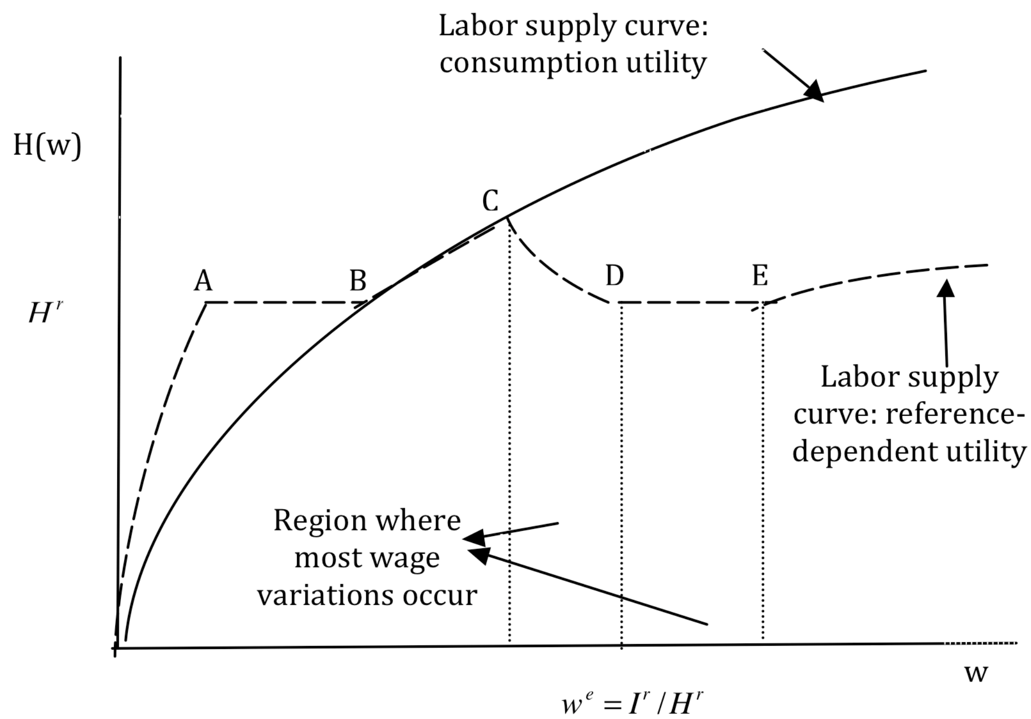


Figure 1.2. A Reference-dependent Driver's Labor Supply Curve

Figure 1.2 compares labor-supply curves for a neoclassical and a reference-dependent driver with the same consumption utility functions. The solid curve is the neoclassical supply curve, and the dashed curve is the reference-dependent one. The shape of the reference- dependent curve depends on which target has a larger influence on the stopping decision, which depends on the relation between the neoclassical optimal stopping point (that is, for consumption utility alone) and the targets. Figure 1.2 illustrates the case suggested by Section 1.3's estimates: For wages that reconcile the income and hours targets as at point D, the neoclassically optimal income and hours are higher than the targets, so the driver stops at his second-reached target. When the wage is to the left of D, the hours target is reached before the income target, and vice versa.

As Figure 1.2 illustrates, reference-dependent labor supply is non-monotonic. When expected earnings per hour is very low, to the left of point A, the higher cost of income losses raises the incentive to work above its neoclassical level (Table 1.1's lower left-hand cell). Along segment AB labor supply is determined by the kink at the hours target, which is reached first. Along segment BC the neoclassical optimal stopping point is above the hours but below the income target, so the gain-loss effects cancel out, and reference-dependent and neoclassical labor supply coincide (Table 1.1's lower right-hand cell). Along segment CD labor supply is determined by the kink at the income target, which is reached second, so that the elasticity of hours with respect to expected earnings per hour is negative. Along segment DE labor supply is determined by the kink at the hours target, which is reached second. (Recall that point D is defined by the wage that is just high enough to reverse which target the driver reaches first.) Finally, when expected earnings per hour is very high, to the right of point E, the higher cost of hours losses lowers the incentive to work below its neoclassical level (Table 1.1's upper right-hand cell). Most realized earnings fall close to point D, either along segment CD where hours decrease with increases in expected earnings per hour because of income targeting, or along segment DE where hours do not change with increases in expected earnings per hour because of hours targeting.[9]

---

[9] There are two possible alternatives to the situation depicted in Figure 1.2. In the first, for earnings that reconcile the income and hours targets, the neoclassical optimal income and hours are lower than the targets, so the driver stops at his first-reached target. This case yields conclusions like Figure 1.2's with

1.3. Econometric Estimates

This section reports econometric estimates of our reference-dependent model of cabdrivers' labor supply. We use Farber's (2005, 2008) data and closely follow his econometric strategies, but instead of treating drivers' targets as latent variables, we treat them as rational expectations and operationalize them via sample proxies with limited endogeneity problems.[10]

Here and in the rest of our econometric analyses, we proxy drivers' point-expectation income and hours targets, driver/day-of-the-week by driver/day-of-the-week, via the analogous sample averages up to but not including the day in question, ignoring sampling variation for simplicity.[11] This avoids confounding from including the current shift's income and hours in the averages, while allowing the targets to vary across days of the week as suggested by the variation of hours and income. This way of proxying the targets loses observations from the first day-of-the-week shift for each

---

some differences in the details. In the second case, the neoclassical optimal income and hours exactly equal the targets, as in KR's preferred personal equilibrium. In that case, near where most realized earnings per hour fall, stopping would be completely determined by the hours target and the income target would have no effect. Thus, our point-expectations version of preferred personal equilibrium is inconsistent with what we find in Farber's data. This does not prove that KR's distributional preferred personal equilibrium would also be inconsistent, but we suspect it would.

[10] Farber generously shared his data with us, now posted at http://www.e-aer.org/data/june08/20030605_data.zip. His 2005 paper gives a detailed description of the data cleaning and relevant statistics. The data are converted from trip sheets recorded by the drivers. These contain information about starting/ending time/location and fare (excluding tips) for each trip. There are in total 21 drivers and 584 trip sheets, from June 2000 to May 2001. Drivers in the sample all lease their cabs weekly so they are free to choose working hours on a daily basis. Because each driver's starting and ending hours vary widely, and 11 of 21 work some night and some day shifts, subleasing seems unlikely. Farber also collected data about weather conditions for control purposes.

[11] There is some risk of bias in ignoring sampling variation, because sampling error tends to be larger early in the sample period. We take this into account by computing estimates with weights equal to the number of realizations (rescaled to sum to the number of observations in each subsample) that are averaged to calculate the expectations. The results are essentially the same as without weighting, with one exception: In Table 1.3's estimates, unweighted estimates would yield an income target that is not significant when we do not distinguish day-of-the-week differences (column 2), but weighted estimates make this parameter significant.

driver because there is no prior information for those shifts.[12] This is a nonnegligible fraction of the total number of observations (3124 out of 13461). But because the criterion for censoring is exogenous and balanced across days of the week and drivers, it should not cause significant bias. When necessary we proxy a driver's expected earnings during the day in the same way, by sample averages, driver/day-of-the-week by driver/day-of-the-week, up to but not including the day in question. This is a noisy proxy, but it is not systematically biased and because it is predetermined it should not cause endogeneity bias.

1.3.1. Probit Models of the Probability of Stopping with a Linear Index Function

We begin by estimating probit models of the probability of stopping with an index function that is linear in cumulative shift hours and cumulative shift income as in Farber (2005), but splitting the sample shift by shift according to whether a driver's earnings for the first x hours of the day (or equivalently, average earnings for the first x hours, but with no need for the average to be constant or independent of history) are higher or lower than his proxied expectations. In estimation we include only observations with cumulative working hours higher than x.

The higher a driver's early earnings, the more likely he is to hit his income

[12] For this reason, we cannot make the sample exactly the same as Farber's, who used only the drivers with a minimum of ten shifts. Strictly speaking, our working hypothesis of rational expectations would justify using averages both prior to and after the shift in question (but still excluding the shift itself). This loses fewer observations, but using only prior sample averages is more plausible and yields somewhat cleaner results. The results are similar using averages after as well as before the shift in question. The average within-driver standard deviation of the income target proxies is $34 and that of the hours target proxies is 1.62 hours. Since for most dates there are only a few driver records, we calculate average across-driver standard deviations day-of-the-week by day-of-the-week, then average across days-of-the-week. The average standard deviation is $37 for the income target proxies and 2.68 hours for the hours target proxies. Thus, the variation across drivers is indeed larger than that within drivers.

target first, simply because early earnings is part of total earnings and can be viewed as a noisy estimate of it. For a wide class of reference-dependent models, including our structural model, a driver's probability of stopping increases at his first-reached target and again (generally by a different amount) at his second-reached target. By contrast, in a neoclassical model, the targets have no effect. This difference is robust to variations in the specification of the targets and the details of the structural specification. Sample-splitting therefore allows a robust assessment of the evidence for reference-dependence, avoiding most of the restrictions needed for structural estimation.

In our model as in Farber's, drivers choose only hours, not effort. Thus early earnings, unlike total earnings, should be approximately uncorrelated with errors in the stopping decision, and so should avoid most problems of sample selection via endogenous variables.

The larger is x the more accurate the split, but we lose the first x hours of observations from each shift, a nonnegligible fraction of the sample if x is large, risking censoring bias. However, if $x = 1$ we lose only 4 shifts (10 trips) out of a total of 584 shifts, so any bias should be small. We report estimates for $x = 1$, but the results are qualitatively robust to values of x up to $x = 5$.[13]

Table 1.2 reports marginal probability effects to maximize comparability with Farber's estimates, but with significance levels computed for the underlying coefficients. In each numbered panel, the left-hand column uses the same specification

[13] When $x > 5$ the sign pattern of estimated coefficients is preserved, but the coefficients are no longer significantly different than 0 in most cases, possibly because of the smaller sample size and censoring bias.

as Farber's (2005) pooled-sample estimates, but with observations deleted as in our split-sample estimates. The center and right-hand columns report our split-sample estimates.

In estimation (1), only income and total hours are used to explain the stopping probability.[14] In the pooled-sample estimates with these controls, coefficients have the expected signs, the effect of hours is significant at the 1% level, and the effect of income is significant at the 10% level. In our split-sample estimates with only these controls, the effect of hours is large and significant whether or not early earnings are higher or lower than expected, but the effect of income is insignificant in either case.

In estimation (2) we control for driver heterogeneity, day-of-the-week, hour of the day, weather, and location. In the pooled sample this yields estimates like those in the left-hand panel, except that the effect of income is now insignificant even at the 10% level. But in the split-sample estimates with this full set of controls, the effect of hours but not that of income is significant at the 1% level when early earnings are higher than expected, while the effect of income is insignificant even at the 10% level; but the effect of income is significant at the 5% level when early earnings are lower than expected, while the effect of hours is insignificant even at the 10% level.

This reversal of the pattern of significant coefficients depending on whether early earnings are higher than expected is inconsistent with a neoclassical model, but is gracefully explained by a reference-dependent model in which stopping probability is usually more strongly influenced by the second target a driver reaches than the first, as in Figure 1.2. Specifically, if the second target reached on a given day normally has

[14] Here we follow Farber (2008) rather than Farber (2005) in using total hours rather than hours broken down into driving hours, waiting hours and break hours, which makes little difference to the results.

Table 1.2. Marginal Effects on the Probability of Stopping: Probit Estimation with Split Samples

	Evaluation Point for Marginal Effect	Pooled data	First hour's earnings> expected	First hour's earnings < expected
Estimation (1)				
Cumulative total hours	8.0	.020*** (.006)	0.022*** (0.006)	0.022*** (0.008)
Cumulative income/100	1.5	0.035* (.016)	0.021 (0.019)	0.021 (0.027)
Other controls		No	No	No
Log likelihood		-1404.905	-688.825	-710.825
Pseudo R2		0.1246	0.1221	0.1333
Estimation (2)				
Cumulative total hours	8.0	0.009*** (0.003)	0.028*** (0.010)	0.005 (0.004)
Cumulative income/100	1.5	0.020 (0.014)	0.035 (0.031)	0.037** (0.025)
Weather (4)		Yes	Yes	Yes
Location (9)		Yes	Yes	Yes
Drivers (21)		Yes	Yes	Yes
Day of week (7)		Yes	Yes	Yes
Hour of day (19)	2:00 p.m.	Yes	Yes	Yes
Log likelihood		-1344.8812	-679.48626	-607.45459
Pseudo R^2		0.2401	0.2550	0.2664
Observation		8958	4664	4294

Note:
Standard errors are computed for the marginal effects to maximize comparability with Farber's estimates, but with significance levels computed for the underlying coefficients rather than the marginal effects: *10%, **5%, ***1%. Robust standard errors clustered by shift are assumed. The subsample estimation weights each observation based on the number of realizations in the history (rescaled to sum to the number of observations in each subsample) used to calculate the proxies for expectations (see footnote 12; results for the unweighted estimation are reported in Appendix 1.1, Table 1.2.A). We use Farber's evaluation point: after 8 total working hours and $150 earnings on a dry day with moderate temperatures in midtown Manhattan at 2:00 p.m. Driver fixed effects and day of week dummies are equally weighted. For dummy variables, the marginal effect is calculated by the difference between values 0 and 1. Following Farber's suggestion, we do not distinguish between driving hours and wait hours between fares. Among the dummy control variables, only driver fixed effects, hour of the day, day of the week, and certain location controls have effects significantly different from 0.

the stronger influence, then on good days, when the income target is reached before the hours target, hours has a stronger influence on stopping probability, as in the *** coefficient in the first row of the right-hand panel of Table 1.2 in the column headed “first hour’s earnings > expected”. On bad days income then has the stronger influence, as in the ** coefficient in the second row of the right-hand panel. By contrast, if the first target reached on a given day usually had the stronger influence, the pattern of significant coefficients would again reverse depending on whether early earnings are higher than expected, but now with a significant influence of income on good days and of hours on bad days. If the population were homogeneous in preferences these would be the only two possible cases, and in that sense the pattern we see is one of two that are characteristic of a reference-dependent model with hours- as well as income-targeting. With heterogeneous preferences other patterns of significance are logically possible, but more “contrived” and so less plausible.

To put these results into perspective, recall that a neoclassical model would predict that hours have an influence on the probability of stopping that varies smoothly with realized income, without regard to whether income is higher than expected. A pure income-targeting model as in Farber (2008) would predict a jump in the probability of stopping when the income target is reached, but an influence of hours that again varies smoothly with realized income. Our estimates are inconsistent with a neoclassical model and—because the effect of hours is significant when income is higher than expected but insignificant when income is lower than expected—with Farber’s income-targeting model. When the utility cost of hours is highly nonlinear, drivers’ neoclassical utility-maximizing choices resemble hours targeting. But neoclassical drivers should still have positive wage elasticity, in contrast to the zero elasticity implied by hours targeting. Further, Section 1.3.3’s structural model can closely approximate a neoclassical model with inelastic labor supply, but there is clear evidence that the hours bunching in the sample follows targets that vary by day-of-the-week in a way that is ruled out by a neoclassical model. Although our estimates are inconsistent with a neoclassical model, they estimates are consistent with our reference-dependent model if the probability of stopping is more strongly influenced by hours when early earnings are higher than expected but by income when lower than expected.

We note again that because the wage elasticity is substantially negative when the income target is the dominant influence on stopping but near zero when the hours target is dominant, the reference-dependent model's distinction between anticipated and unanticipated wage changes can reconcile an anticipated wage increase's positive incentive to work with a negative aggregate wage elasticity of hours. Finally, with a distribution of realized wages, the model can also reproduce Farber's (2005) findings that aggregate stopping probabilities are significantly related to hours but not realized earnings, and that they respond smoothly to earnings.

1.3.2. Reduced-form Estimates of the Probability of Stopping

We now estimate a reduced-form model of stopping probability, with dummy variables to measure the increments due to hitting the income and hours targets as in Farber's (2008) Table 2, but with the sample proxies for targets introduced above instead of Farber's estimated targets.

Table 1.3 reports reduced-form estimates of the increments in stopping probability on hitting the estimated income and hours targets. The estimated coefficients of dummy variables indicating whether earnings or hours exceeds the targets are positive, the sign predicted by a reference-dependent model, and significantly different from 0. The estimates confirm and extend the results from our split-sample probits, in that the significant effects of income and hours come mainly from whether they are above or below their targets rather than from their levels. The level of income has a slightly negative, insignificant effect and the level of hours has a positive, significant effect. In this respect the estimates suggest that hours have a nonnegligible neoclassical effect as well as their reference-dependent effect.

Table 1.3. Marginal Effects on the Probability of Stopping: Reduced-Form Model Allowing Jumps at the Targets

	Using driver specific sample average income and hours prior to the current shift as targets		Using driver and day-of-the-week specific sample average income and hours prior to the current shift as targets	
	(1)	(2)	(3)	(4)
Cumulative total hours>hours target	0.036***	0.030***	0.055***	0.155***
	(0.013)	(0.022)	(0.016)	(0.051)
Cumulative income > income target	0.058***	0.020*	0.049***	0.056**
	(0.018)	(0.017)	(0.017)	(0.036)
Cumulative total hours	0.011***	0.007**	0.012***	0.018**
	(0.005)	(0.006)	(0.004)	(0.010)
Cumulative Income/100	-0.010	0.010	-0.012	0.016
	(0.015)	(0.016)	(0.013)	(0.039)
Weather (4)	No	Yes	No	Yes
Locations (9)	No	Yes	No	Yes
Drivers (21)	No	Yes	No	Yes
Days of the week (7)	No	Yes	No	Yes
Hour of the day (19)	No	Yes	No	Yes
Log likelihood	-1526.9354	-1367.8075	-1493.3419	-1349.809
Pseudo R^2	0.1597	0.2472	0.1756	0.2740
Observation	10337	10337	10337	10337

Note:
Standard errors are computed for the marginal effects to maximize comparability with Farber’s estimates, but with significance levels computed for the underlying coefficients rather than the marginal effects: *10%, **5%, ***1%. Robust standard errors clustered by shift are assumed. The estimation weights each observation based on the number of realizations in the history (rescaled to sum to the number of observations in each estimation) used to calculate proxies for expectations (see footnote 12; results for the unweighted estimation are reported in online Appendix 1.1, Table 1.2.A). We use Farber’s evaluation point: after 8 total working hours and $150 earnings on a dry day with moderate temperatures in midtown Manhattan at 2:00 p.m. Driver fixed effects and day of week dummies are equally weighted. For dummy variables, the marginal effect is calculated by the difference between values 0 and 1. As in Farber (2008) (but no Farber (2005), we do not distinguish between driving hours and waiting hours between fares. Among the dummy control variables, only driver fixed effects, hour of the day, day of the week, and certain location controls have effects significantly different from 0.

1.3.3. Structural Estimation

We now estimate Section 1.2's structural model. Our structural model makes no sharp general predictions. In particular, whether the aggregate stopping probability is more strongly influenced by income or hours depends on estimated parameters and how many shifts have realized income higher than expected. Even so, structural estimation is an important check on the model's ability to give a useful account of drivers' labor supply.

We use the same sample proxies for drivers' targets as before, and we take a driver's expectations about earnings during the day as predetermined rational expectations, proxied by sample averages, driver/day-of-the-week by driver/day-of-the-week, up to but not including the day in question. This proxy is noisy, but it is not a source of endogeneity or other bias.

Section 1.2 explains the model. In the structural estimation, as in Farber (2008), we impose the further assumption that consumption utility has the functional form $U(I,H) = I - \frac{\theta}{1+\rho} H^{1+\rho}$, where ρ is the elasticity of the marginal rate of substitution. Thus, the driver has constant marginal utility of income (and is risk-neutral in it, treating $U(\cdot)$ as a von Neumann-Morgenstern utility function), in keeping with the fact that income is storable and the day is a small part of his economic life. However, he is averse to hours as in a standard labor supply model.

Substituting this functional form into (1.1)-(1.2) yields:

$$(1.3)\, V(I,H \mid I^r,H^r) = (1-\eta)\left[I - \frac{\theta}{1+\rho}H^{1+\rho}\right] + \eta\left[1_{(I-I^r\le 0)}\lambda(I-I^r) + 1_{(I-I^r>0)}(I-I^r)\right]$$

$$-\eta\left[1_{(H-H^r\ge 0)}\lambda\left[\frac{\theta}{1+\rho}H^{1+\rho} - \frac{\theta}{1+\rho}(H^r)^{1+\rho}\right]\right] - \eta\left[1_{(H-H^r<0)}\left[\frac{\theta}{1+\rho}H^{1+\rho} - \frac{\theta}{1+\rho}(H^r)^{1+\rho}\right]\right].$$

Like Farber, we assume that the driver decides to stop at the end of a given trip if and only if his anticipated gain in utility from continuing work for one more trip is negative. Again letting I_t and H_t denote income earned and hours worked by the end of trip t, this requires:

$$(1.4) \qquad E[V(I_{t+1}, H_{t+1}|I^r,H^r)] - V(I_t, H_t|I^r,H^r) + x_t\beta + c + \varepsilon < 0,$$

where $I_{t+1} = I_t + E(f_{t+1})$ and $H_{t+1} = H_t + E(h_{t+1})$, and $E(h_{t+1})$ are the next trip's expected fare and time (searching and driving), $x_t\beta$ include the effect of control variables, c is the constant term, and ε is a normal error with mean zero and variance σ^2. We estimate a non-zero constant term to avoid bias, even though theory suggests $c = 0$.

Appendix 1.2 gives the details of deriving the likelihood function

(1.5)

$$\sum_{i=1}^{584}\sum_{t=i}^{T_i} \ln\Phi[((1-\eta+\eta\lambda)a_{1,it} + a_{2,it} - (1-\eta+\eta\lambda)\frac{\theta}{\rho+1}b_{1,it}(\rho) - \frac{\theta}{\rho+1}b_{2,it}(\rho) + x_t\beta + c)/\sigma]$$

, where i refers to the shift and t to the trip within a given shift, and $a_{1,it}, a_{2,it}, b_{1,it}(\rho)$, and $b_{2,it}(\rho)$ are shorthands for components of the right-hand side of (1.3), as explained in Appendix 1.2.

Here, unlike in a standard probit model, σ is identified through $a_{2,it}$, which represents the change in income "gain" relative to the income target. However, as is clear from the likelihood function, η and λ cannot be separately identified: Only $1 - \eta + \eta\lambda$, the factor by which (directly or inversely) the reference-dependent marginal rate of substitution differs from the neoclassical marginal rate of substitution (Table 1.1) is identified. If $1 - \eta + \eta\lambda = 1$, or equivalently $\eta(\lambda - 1) = 0$, the model reduces to a neoclassical model. This happens trivially if $\eta = 0$ so there is no weight on gain-loss utility, or if $\eta \neq 0$ but $\lambda = 1$ so gains and losses are weighted equally. If $\eta = 1$ the model has only gain-loss utility as was usually assumed before KR (2006), and $1 - \eta + \eta\lambda = \lambda$. In that sense our estimates of $1 - \eta + \eta\lambda$ are directly comparable to most estimates of the coefficient of loss aversion that have been reported in the literature.

Table 1.4 reports structural estimates, expanded to identify the effects of different proxies and the reasons for the differences between our and Farber's (2008) results, and to allow different coefficients of loss aversion, λ_H and λ_I, for hours and income. Column 1's baseline model yields plausible parameter estimates that confirm and refine the conclusions of Section 1.3.1-2's analyses. For both λ_H *and* λ_I, the null hypothesis that $\eta(\lambda - 1) = 0$ is rejected at the 1% level, ruling out the restrictions $\eta = 0$ or $\lambda = 1$that would reduce the model to a neoclassical model.[15] For both λ_H *and* λ_I, the implied estimate of $1 - \eta + \eta\lambda$ ($= 1 + \eta(\lambda - 1)$) is comparable to most reported estimates

[15] The estimated standard errors suggest that $\eta(\lambda-1)$ is not significantly different from zero in most specifications, based on the Wald Test. Here we use likelihood ratio tests, which give results somewhat different from the Wald Test. There are at least two reasons why the likelihood ratio test might give different results: First, some parameter transformations are needed to facilitate numerical estimation, and the likelihood ratio test is invariant to such transformations under maximum likelihood estimation, but the Wald Test is not invariant. Second, although both test statistics converge to the Chi-square distribution asymptotically, for small samples the likelihood ratio test statistic is closer to the Chi-square distribution used for inference. Because our sample size is quite large, the first reason is probably the more important one.

of the coefficient of loss aversion. The hypothesis that $\lambda_H = \lambda_I$ cannot be rejected, although the estimated λ_H robustly exceeds λ_I.

Table 1.4. Structural Estimates under Alternative Specifications of Expectations

	(1)	(2)	(3)	(4)	(5)
$\eta(\lambda_H - 1)$	1.309***	1.886***	0.671***	0.188***	-
[p-value]	[0.000]	[0.000]	[0.000]	[0.001]	
$\eta(\lambda_I - 1)$	0.512***	0.299**	0.256***	0.111*	2.007***
[p-value]	[0.001]	[0.041]	[0.002]	[0.057]	[0.000]
θ	0.035***	0.017***	0.043***	0.152***	0.018***
[p-value]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
ρ	0.597***	0.782***	0.566***	0.212***	1.407***
[p-value]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
σ^+	0.127	0.117	0.072	0.045	0.286
[p-value]	[0.253]	[0.104]	[0.996]	[0.280]	[0.484]
c	-0.047	0.014	-0.045	0.029	-0.036
[p-value]	[0.710]	[0.929]	[0.825]	[0.755]	[0.902]
Test $\lambda_H = \lambda_I$ [p-value]	[0.243]	[0.112]	[0.997]	[0.666]	-
Observations	10337	10337	10337	10337	10337
Log-likelihood	-1321.1217	-1326.3005	-1312.8993	-1367.2374	-1333.0964

Notes:
1. Column (1): Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and the next-trip earnings/times expectation. Column (2): Use driver and day-of-the-week specific sample averages prior and after the current shift as the income/hours targets and next-trip the earnings/times expectation. Column (3): Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and fit the sophisticated next-trip earnings/time expectation. Column (4): Use driver (without day-of-the-week difference) specific sample averages prior to the current shift as income/hours targets and the next-trip earnings/time expectation. Column (5): Income target only: use driver and day-of-the-week specific sample averages prior to the current shift as income target and next-trip earnings/time expectation
2. Significance levels *10%, **5%, ***1%. We perform likelihood ratio tests on each estimated parameter and indicate the corresponding p-values and significance levels. The null-hypothesis is that each parameter equals zero except for the variance estimate where we test $\sigma = 1$. The estimation weights each observation based on the number of realizations in the history (rescaled to sum to the number of observations in each estimation) used to calculate proxies for expectations (see footnote 12; results for the unweighted estimation are reported in Appendix 1.1, Table 1.3.A). Control variables include driver fixed effects (18), day of week (6), hour of day (18), location(8), and weather (4).

Columns 2-5 change one thing at a time from the baseline. Column 2 confirms the robustness of Column 1's results to basing targets on sample proxies after as well as before the current shift (but still omitting the current shift; see footnote 13). Column 3 confirms the robustness of Column 1's results to more sophisticated earnings forecasting, via a model of next-trip fare/time expectations using the 3124 observations omitted from the first shifts for each day-of-the-week for each driver, and estimated using the current sample.[16] Column 4 suggests that Column 1's results are *not* robust to ruling out day-of-the-week differences as in Farber (2008): This restriction obscures the effects of reference-dependence, in that the effects of the targets become smaller and in one case significant only at the 10% level. By contrast, Column 5 suggest that Column 1's results are robust to Farber's (2008) restriction to income- but not hours-targeting.

Table 1.4's five models all have the same number of parameters except for column 5, which has no loss aversion coefficient for the hours target: a constant term, five structural parameters, and 55 controls. Our proxies for targets and trip-level expectations are either calculated as sample averages or as predicted values with coefficients estimated out of sample, and this choice does not affect the number of parameters. Although Farber (2008) argues that a reference-dependent model has too many degrees of freedom to be fairly compared with a neoclassical model —a loss aversion coefficient and heterogeneous income targets—defining the targets as rational expectations reduces the difference.

Column 3's model, with drivers sophisticated enough to predict future wages based on location, clock hours, etc., fits best. Of the remaining four models, all with

[16] The other variables include day-of-the-week, hour-of-the-day, locations at the end of the trip, and weather controls. Surprisingly, there is not much variation by time of day, but there is a lot of variation across locations. Appendix 1.3, Table 1.6 reports the trip fares and time estimates whose fitted values are used to proxy drivers' expectations in those models.

Table 1.5. Estimated Optimal Stopping Times (in Hours)

Hourly earnings	(1)		(2)		(3)		(4)	
	A	B	A	B	A	B	A	B
	θ =0.006 ρ =1.494	θ= 0.035 ρ= 0.597 $\eta(\lambda_H - 1)$ =1.309 $\eta(\lambda_I - 1)$ = 0. 512	θ =0.002 ρ =1.942	θ=0.017 ρ=0.782 $\eta(\lambda_H - 1)$ =1.886 $\eta(\lambda_I - 1)$ = 0.299	θ =0.089 ρ =1.210	θ =0.043 ρ= 0.566 $\eta(\lambda_H - 1)$ =0.67 $\eta(\lambda_I - 1)$ = 0.256	θ =0.030 ρ =0.901	θ = 0.152 ρ = 0.212 $\eta(\lambda_H - 1)$ =0.188 $\eta(\lambda_I - 1)$ =0.111
\$16.7	9.27	9.58^{I}	9.76	9.58^{I}	1.68	9.58^{I}	6.72	3.51
\$18.4	9.89	8.70^{I}	10.26	8.70^{I}	1.82	8.70^{I}	7.49	5.55
\$19.4	10.24	8.25^{I}	10.55	8.25^{I}	1.90	8.25^{I}	7.94	7.12
\$20.3	10.56	7.88^{I}	10.79	7.88^{I}	1.98	7.88^{I}	8.35	7.88^{I}
\$21.3	10.91	7.80^{H}	11.06	7.80^{H}	2.06	7.80^{H}	8.81	7.80^{H}
\$22.0	11.14	7.80^{H}	11.25	7.80^{H}	2.11	7.80^{H}	9.13	7.80^{H}
\$22.8	11.41	7.80^{H}	11.46	7.80^{H}	2.18	7.80^{H}	9.50	7.80^{H}
\$23.8	11.75	7.80^{H}	11.72	7.80^{H}	2.25	8.30	9.96	7.80^{H}
\$25.3	12.24	7.80^{H}	12.09	8.10	2.37	9.20	10.66	7.80^{H}
Correlation of hourly earnings and working hours	0.99	-0.83	0.99	-0.75	0.99	-0.27	0.97	0.80

Note:
1. Column (1): Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and the next-trip earnings/times expectation. Column (2): Use driver and day-of-the-week specific sample averages prior and after the current shift as the income/hours targets and next-trip the earnings/times expectation. Column (3): Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and fit the sophisticated next-trip earnings/time expectation. Column (4): Use driver (without day-of-the-week difference) specific sample averages prior to the current shift as income/hours targets and the next-trip earnings/time expectation.
2. The chosen wages are 10-90 percentiles of the wage distribution in the sample. For illustrative purposes we take the average income (\$160) and working hours (7.8) in the estimation sample as income and hours targets to determine the optimal working hours given the estimated coefficients. For each model, we calculate both the neoclassical optimal working hours (column A) based on a separate estimation that includes consumption utility only, and the reference-dependent optimal working hours (column B) based on estimates from table 4 that include both consumption utility and the gain-loss utility. Optimal working hours superscripted H or I denotes that the number is bounded by the hours or income target.

constant expectations throughout the shift, Column 1's model, the baseline, fits best.

The likelihood cost of ruling out sophisticated earnings forecasting is nontrivial, though this does not seem to distort the parameter estimates much. Despite Column 5's robustness result, the likelihood cost of ruling out hours-targeting is also nontrivial, as is that of ruling out day-of-the-week differences.

To illustrate the implications of the estimated utility function parameters under Table 1.4's alternative specifications, Table 1.5 presents the optimal stopping times implied by our estimates of the structural reference-dependent model for each specification (column B) and for representative percentiles of the observed distribution of realized wages, with "neoclassical" optimal working hours (column A) for comparison, computed from the estimates of a separate structural neoclassical model with consumption utility only.[17] The implied reference-dependent stopping times seem reasonable for all four models. However, for the model of Column 3 neoclassical working hours are very low, perhaps because the model is estimated with varying next-trip earnings and time predictions and so simulating the optimal working hours with a constant wage is inappropriate. By contrast, neoclassical working hours of other columns are in a reasonable range. For the model of Column 4 the neoclassical optimal solution ranges from below to above the targets as earnings per hour vary, so in a reference-dependent model labor supply is driven by neoclassical considerations for low earnings but by the hours target for high earnings; in aggregate the correlation between earnings per hour and optimal working hours is positive.

[17] In Column A we compute optimal working hours after reestimating the parameters of the consumption utility function constraining $\eta = 1$. Appendix 1.4, Table 1.7 gives the implied average stopping probabilities for various ranges relative to the targets. Our estimates imply comparatively little bunching around the targets. Even so, the targets have a very strong influence on the stopping probabilities, and the second-reached target has a stronger effect than the first-reached target.

Like Section 1.3.1's probits, our structural model resolves the apparent contradiction between a negative aggregate wage elasticity and the positive incentive to work of an anticipated increase in expected earnings per hour. In our model the stopping decisions of some drivers, on some days, will be more heavily influenced by their income targets, in which case their earnings elasticities will be negative, while the decisions of other drivers on other days will be more heavily influenced by their hours targets, with elasticities close to zero. When $\eta(\lambda - 1)$ is large enough, and with a significant number of observations in the former regime, the model will yield a negative aggregate elasticity. To illustrate, Table 1.5 also reports each specification's implication for the aggregate correlation of earnings and optimal working hours, a proxy for the elasticity. All reference-dependent models but column (4), which suppresses day-of-the-week differences, have a negative correlation between earnings per hour and optimal working hours.

Despite the influence of the targets on stopping probabilities, the heterogeneity of realized earnings yields a smooth aggregate relationship between stopping probability and realized income, so the model can reconcile Farber's (2005) finding that aggregate stopping probabilities are significantly related to hours but not income with a negative aggregate wage elasticity of hours as found by Camerer et al. (1997).

Finally, our structural model avoids Farber's (2008) criticism that drivers' estimated targets are too unstable and imprecisely estimated to allow a useful reference-dependent model of labor supply. The key function $\eta(\lambda - 1)$ of the parameters of gain-loss utility is plausibly and precisely estimated, robust to the specification of proxies for drivers' expectations, and comfortably within the range that indicates reference-dependent preferences.

1.4. Conclusion

In this paper we have proposed and estimated a model of cabdrivers' labor supply based on KR's theory of reference-dependent preferences, with targets for hours as well as income, both determined by proxied rational expectations. Our analysis builds on Farber's (2005, 2008) empirical analyses, which allowed income- but not hours-targeting and treated the targets as latent variables.

Our model, estimated with Farber's data, suggests that reference-dependence is an important part of the labor-supply story in his dataset, and that using KR's model to take it into account does yield a useful model of cabdrivers' labor supply. Overall, our results suggest that a more comprehensive investigation of the behavior of cabdrivers and other workers with similar choice sets, with larger datasets and more careful modeling of targets, will yield a reference- dependent model of labor supply that significantly improves upon the neoclassical model.

Acknowledgements

Chapter 1, in full, is currently being prepared for publication as a forthcoming paper in *American Economic Review.* Crawford, Vincent P.; Meng, Juanjuan. The dissertation author was one of the primary authors of this paper.

Appendix 1.1. Estimations of Tables 1.2-1.4 Without Weights

Table 1.2.A. Probability of Stopping: Linear Probits with Split Sample

	Evaluation Point for Marginal Effect	Pooled data	First hour's earnings> expected	First hour's earnings < expected
Estimation (1)				
Cumulative total hours	8.0	.020*** (.006)	0.025*** (0.006)	0.017** (0.009)
Cumulative income/100	1.5	0.035* (.016)	0.030 (0.020)	0.037 (0.026)
Other controls		No	No	No
Log likelihood		-1550.452	-806.30573	-742.87617
Pseudo R2		0.1239	0.1314	0.1172
Estimation (2)				
Cumulative total hours	8.0	0.009*** (0.003)	0.026*** (0.009)	0.004 (0.004)
Cumulative income/100	1.5	0.020 (0.014)	0.034 (0.028)	0.029* (0.022)
Weather (4)		Yes	Yes	Yes
Location (9)		Yes	Yes	Yes
Drivers (21)		Yes	Yes	Yes
Day of week (7)		Yes	Yes	Yes
Hour of day (19)	2:00 p.m.	Yes	Yes	Yes
Log likelihood		-1344.8812	-683.58849	-628.45562
Pseudo R^2		0.2401	0.2636	0.2532
Observation		8958	4664	4294

Note:
Standard errors are computed for the marginal effects to maximize comparability with Farber's estimates, but with significance levels computed for the underlying coefficients rather than the marginal effects: *10%, **5%, ***1%. Robust standard errors clustered by shift are assumed. We use Farber's evaluation point: after 8 total working hours and $150 earnings on a dry day with moderate temperatures in midtown Manhattan at 2:00 p.m. Driver fixed effects and day of week dummies are equally weighted. For dummy variables, the marginal effect is calculated by the difference between values 0 and 1. As in Farber (2008) (but no Farber (2005), we do not distinguish between driving hours and waiting hours between fares. Among the dummy control variables, only driver fixed effects, hour of the day, day of the week, and certain location controls have effects significantly different from 0.

Table 1.3.A. Marginal Effects on the Probability of Stopping:
Reduced-Form Model Allowing Jumps at the Targets

	Using driver specific sample average income and hours prior to the current shift as targets		Using driver and day-of-the-week specific sample average income and hours prior to the current shift as targets	
	(1)	(2)	(3)	(4)
Cumulative total hours>hours target	0.040***	0.065***	0.047***	0.109***
	(0.013)	(0.031)	(0.014)	(0.039)
Cumulative income > income target	0.052***	0.024	0.043***	0.038*
	(0.06)	(0.025)	(0.015)	(0.024)
Cumulative total hours	0.012***	0.019***	0.013***	0.016***
	(0.004)	(0.009)	(0.004)	(0.008)
Cumulative Income/100	-0.007	0.024	-0.001	0.006
	(0.013)	(0.035)	(0.015)	(0.029)
Weather (4)	No	Yes	No	Yes
Locations (9)	No	Yes	No	Yes
Drivers (21)	No	Yes	No	Yes
Days of the week (7)	No	Yes	No	Yes
Hour of the day (19)	No	Yes	No	Yes
Log likelihood	-1546.1866	-1369.5477	-1535.036	-1349.809
Pseudo R^2	0.1630	0.2587	0.1691	0.2693
Observation	10337	10337	10337	10337

Note:
Standard errors are computed for the marginal effects to maximize comparability with Farber's estimates, but with significance levels computed for the underlying coefficients rather than the marginal effects: *10%, **5%, ***1%. Robust standard errors clustered by shift are assumed. We use Farber's evaluation point: after 8 total working hours and $150 earnings on a dry day with moderate temperatures in midtown Manhattan at 2:00 p.m. Driver fixed effects and day of week dummies are equally weighted. For dummy variables, the marginal effect is calculated by the difference between values 0 and 1. As in Farber (2008) (but no Farber (2005), we do not distinguish between driving hours and waiting hours between fares. Among the dummy control variables, only driver fixed effects, hour of the day, day of the week, and certain location controls have effects significantly different from 0.

Table 1.4.A. Structural Estimates under Alternative Specifications of Expectations

	(1)	(2)	(3)	(4)	(5)
$\eta(\lambda_H - 1)$	2.338***	4.327***	0.872***	0.237***	-
[p-value]	[0.000]	[0.000]	[0.000]	[0.000]	
$\eta(\lambda_I - 1)$	0.631***	0.610***	0.267***	0.044*	3.163***
[p-value]	[0.004]	[0.000]	[0.008]	[0.0594]	[0.000]
θ	0.015***	0.020***	0.018***	0.099***	0.014***
[p-value]	[0.000]	[0.000]	[0.000]	[0.000]	[.000]
ρ	0.839***	0.483	0.883***	0.258***	1.645***
[p-value]	[0.003]	[0.403]	[0.000]	[0.000]	[0.000]
σ^+	0.196	0.185	0.096	0.040	0.539***
[p-value]	[0.168]	[0.293]	[0.996]	[0.105]	[0.757]
c	0.007	0.006	-0.012	0.134	0.138
[p-value]	[0.954]	[0.958]	[0.998]	[0.782]	[0.719]
Test $\lambda_H = \lambda_I$ [p-value]	[0.214]	[0.177]	[0.996]	[0.204]	-
Observations	10337	10337	10337	10337	10337
Log-likelihood	-1360.9672	-1361.711	-1351.4242	-1368.8756	-1371.8068

Notes:

1. Column (1): Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and the next-trip earnings/times expectation. Column (2): Use driver and day-of-the-week specific sample averages prior and after the current shift as the income/hours targets and next-trip the earnings/times expectation. Column (3): Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and fit the sophisticated next-trip earnings/time expectation. Column (4): Use driver (without day-of-the-week difference) specific sample averages prior to the current shift as income/hours targets and the next-trip earnings/time expectation. Column (5): Income target only: use driver and day-of-the-week specific sample averages prior to the current shift as income target and next-trip earnings/time expectation.
2. Significance levels *10%, **5%, ***1%. We perform likelihood ratio tests on each estimated parameter and indicate the corresponding *p*-values and significance levels. The null-hypothesis is that each parameter equals zero except for the variance estimate where we test $\sigma = 1$. The estimation weights each observation based on the number of realizations in the history (rescaled to sum to the number of observations in each estimation) used to calculate proxies for expectations (see footnote 12; results for the unweighted estimation are reported in online Appendix A, Table A3). Control variables include driver fixed effects (18), day of week (6), hour of day (18), location(8), and weather (4).

Appendix 1.2. Derivation of the Likelihood Function in the Structural Estimation

Recall that given equation (1.3) and (1.4), the likelihood function can now be written, with *i* denoting the shift and *t* the trip within a given shift, as:

(1.6)
$$\sum_{i=1}^{584}\sum_{t=i}^{T_i} \ln\Phi[((1-\eta)(A_{it}-\frac{\theta}{\rho+1}B_{it}(\rho))+\eta(\lambda a_{1,it}+a_{2,it})$$
$$-\lambda\frac{\theta}{\rho+1}b_{1,it}(\rho)-\frac{\theta}{\rho+1}b_{2,it}(\rho))+x_t\beta+c)/\sigma].$$

Where

$A_{it}=I_{i,t+1}-I_{i,t}$.

$B_{it}(\rho)=H_{i,t+1}^{\rho+1}-H_{i,t}^{\rho+1}$.

$a_{1,it}=1_{(I_{i,t+1}-I_i^r\le 0)}(I_{i,t+1}-I_i^r)-1_{(I_{i,t}-I_i^r\le 0)}(I_{i,t}-I_i^r)$.

$a_{2,it}=1_{(I_{i,t+1}-I_i^r>0)}(I_{i,t+1}-I_i^r)-1_{(I_{i,t}-I_i^r>0)}(I_{i,t}-I_i^r)$.

$b_{1,it}(\rho)=1_{(H_{i,t+1}-H_i^r\ge 0)}(H_{i,t+1}^{\rho+1}-(H_i^r)^{\rho+1})-1_{(H_{i,t}-H_i^r\ge 0)}(H_{i,t}^{\rho+1}-(H_i^r)^{\rho+1})$.

$b_{2,it}(\rho)=1_{(H_{i,t+1}-H_i^r<0)}(H_{i,t+1}^{\rho+1}-(H_i^r)^{\rho+1})-1_{(H_{i,t}-H_i^r<0)}(H_{i,t}^{\rho+1}-(H_i^r)^{\rho+1})$.

Note that we have $A_{it}=a_{1,it}+a_{2,it}$ and $B_{it}=b_{1,it}(\rho)+b_{2,it}(\rho)$.

Substituting these equations yields a reduced form for the likelihood function as expressed by equation (1.5).

Appendix 1.3. Trip Fares and Time Estimates Whose Fitted Values are Used as Proxies for Drivers' Expectations in Table 1.4, column 3

Table 1.6. Trip Fares and Time Estimates Whose Fitted Values Are Used as Proxies for Drivers' Sophisticated Expectations in Table 1.4

	Time	Fare		Time	Fare
Clock hours			Day of the Week		
0	-0.100	0.006	Monday	0.017	-
	(0.228)	(0.022)		(0.025)	-
1	-0.121	-0.005	Tuesday	-0.007	0.001
	(0.231)	(0.022)		(0.023)	(0.003)
2	-0.255	-0.025	Wednesday	-0.012	-0.002
	(0.239)	(0.024)		(0.023)	(0.004)
3	-0.193	-	Thursday	0.013	0.004
	(0.265)	-		(0.023)	(0.004)
4	-	0.026	Friday	-0.003	-0.000
	-	(0.039)		(0.023)	(0.003)
5 - 10	-0.022	-0.006	Saturday	0.038*	0.006*
	(0.226)	(0.021)		(0.022)	(0.003)
11	-0.022	-0.011	Sunday	-	0.001
	(0.227)	(0.022)		-	(0.004)
12	0.026	-0.005	Mini temp < 30	0.016	0.000
	(0.227)	(0.022)		(0.027)	(0.004)
13	-0.032	-0.001	Max temp > 80	0.019	-0.002
	(0.227)	(0.021)		(0.023)	(0.003)
14	-0.074	-0.003	Hourly rain	-0.147	-0.073
	(0.227)	(0.021)		(0.317)	(0.046)
15	-0.084	-0.005	Daily snow	0.006	0.000
	(0.227)	(0.021)		(0.010)	(0.001)
16	-0.074	0.007	Downtown	-0.025	0.013
	(0.227)	(0.022)		(0.121)	(0.018)
17	-0.132	-0.006	Midtown	-0.066	0.001
	(0.226)	(0.021)		(0.120)	(0.018)
18	-0.152	-0.010	Uptown	-0.036	0.003
	(0.226)	(0.021)		(0.121)	(0.018)
19	-0.189	-0.016	Bronx	-	-
	(0.226)	(0.021)		-	-
20	-0.137	-0.006	Queens	0.337**	0.080***
	(0.226)	(0.021)		(0.151)	(0.022)
21	-0.160	-0.008	Brooklyn	0.180	0.052***
	(0.226)	(0.021)		(0.135)	(0.020)
22	-0.177	-0.004	Kennedy Airport	0.645***	0.164***
	(0.226)	(0.021)		(0.136)	(0.020)
23	-0.128	0.003	LaGuardia Airport	0.333**	0.110***
	(0.226)	(0.021)		(0.130)	(0.019)
Constant	0.307	0.051*	Others	0.154	0.030
	(0.260)	(0.029)		(0.156)	(0.023)
Driver dummy 21	Yes	Yes	R2	0.122	0.202
Observations	2989	2989		2989	2989

Appendix 1.4. Implied Average Probabilities of Stopping for Various Ranges

Table 1.7. Implied Average Probabilities of Stopping for Various Ranges Relative to the Targets

	(1)	(2)	(3)	(4)
Wage in the first hour > expected				
Before income target	0.020	0.021	0.019	0.022
At income target	0.083	0.097	0.080	0.092
In between two targets	0.105	0.109	0.103	0.103
At hours target	0.159	0.148	0.139	0.134
Above hours target	0.175	0.156	0.175	0.150
Wage in the first hour < expected				
Before hours target	0.0180	0.0193	0.018	0.021
At hours target	0.081	0.086	0.094	0.094
In between two targets	0.106	0.109	0.113	0.119
At income target	0.161	0.148	0.181	0.138
Above income target	0.188	0.180	0.187	0.164

Note:
1. Column1: Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and the next-trip earnings/times expectation. Column2: Use driver and day-of-the-week specific sample averages prior and after the current shift as the income/hours targets and next-trip the earnings/times expectation. Column3: Use driver and day-of-the-week specific sample averages prior to the current shift as the income/hours targets and fit the sophisticated next-trip earnings/time expectation. Column4: Use driver (without day-of-the-week difference) specific sample averages prior to the current shift as income/hours targets and the next-trip earnings/time expectation.
2. The probability of each range is calculated from the average predicted probabilities of trips. A range is two-sided with tolerance 0.1: before target means < 0.95×target; at target means > 0.95×target but < 1.05×target; and above target means > 1.05×target. The probabilities are first computed for each driver and range and then averaged across drivers within each range, hence do not sum to one.

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review*, 101, in press.

Camerer, Colin, Linda Babcock, George Loewenstein and Richard Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics*, 112(2): 407-441.

Crawford, Vincent P., and Juanjuan Meng. 2008. "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." U. C. S. D. Discussion Paper 2008-03, http://repositories.cdlib.org/ucsdecon/2008-03/.

Doran, Kirk. 2009. "Reference Points, Expectations, and Heterogeneous Daily Labor Supply." Manuscript, University of Notre Dame.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263-292.

Farber, Henry S. 2005. "Is Tomorrow another Day? The Labor Supply of New York City Cabdrivers." *Journal of Political Economy*, 113(1): 46-82.

Farber, Henry S. 2008. "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review*, 98(2): 1069-1082.

Fehr, Ernst, and Lorenz Goette. 2007. "Do Workers Work More if Wages are Higher? Evidence from a Randomized Field Experiment." *American Economic Review*, 97(1): 298-317.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133-1165.

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047-1073.

Kőszegi, Botond and Matthew Rabin. 2009. "Reference-Dependent Consumption Plans," *American Economic Review*, 99(3): 909-936.

Oettinger, Gerald S. 1999. "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors." *Journal of Political Economy*, 107(2): 360-392.

Post, Thierry, Martijn J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game show." *American Economic Review*, 98(1): 38-71.

Tversky, Amos, and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." Quarterly Journal of Economics, 106(4): 1039-1061.

Chapter 2

The Disposition Effect and Expectations as Reference Point

2.1. Introduction

The disposition effect refers to the observation that stock market investors tend to hold on to their losers for too long and sell their winners too soon, with losers and winners defined by comparing current price to the initial or (when shares were acquired at different times) the average purchase price (Shefrin and Statman (1985), Odean (1998), and Weber and Camerer (1998)).[18] Odean (1998), for instance, analyzes trading records of individual investors at a large discount brokerage house. He finds a strong asymmetry in the sale probabilities of stocks that currently show a gain and those that show a loss relative to the average purchase price.

In Odean's (1998) dataset, most investors do not immediately purchase another stock after selling an old one, so the selling decision is largely a choice between holding a risky stock or safe cash, which mainly reflects attitudes toward risk.[19] The

[18] For consistency with the previous literature, I use this definition of winners and losers below when talking about patterns in the data, even though the true reference level of price may not be the initial or average purchase price. Further, Odean's (1998) analysis suggests that the choice whether to use the initial or the average purchase price makes little difference empirically, so I focus below on average purchase for simplicity.

[19] It is of course always possible that investors purchase other types of risky assets, or they have accounts in other brokerage companies so that the trading records in this sample are not complete. However, given the large number of investors and the long period involved, the time lag between each sale and the next purchase should largely reflect an investment pattern rather than these incidences. For instance, it is not likely that most investors trade in another asset market every time they sell a stock; or they use stock accounts in other brokerage companies to buy a new stock when they sell an old one.

disposition effect thus poses a challenge to explanations based on simple models with expected-utility maximizing investors, in that there is no reason why the sharp changes in risk aversion needed to explain the disposition effect should bunch around the average purchase price, especially when investors have varying wealth levels, different starting portfolios, and distinct purchase prices. Further, Odean explicitly considers expected-utility explanations for the asymmetry in sale probabilities based on richer specifications of the investor's problem, finding that portfolio rebalancing, transaction costs, taxes, or rationally anticipated mean reversion cannot explain the observed asymmetry.[20]

The most popular informal explanation of the disposition effect has been prospect theory (Kahneman and Tversky (1979), and Odean (1998)). Prospect theory assumes that the carrier of utility is not the level of wealth, but the change in wealth relative to a reference point. The theory has three main elements: Loss aversion (losses relative to the reference point hurt investors more than gains please them), diminishing sensitivity (investors are less sensitive to big gains and losses than small ones), and nonlinear probability weighting (investors systematically overweight small probabilities). [21]

---

[20] Weber and Camerer (1998) find that incorrect beliefs of mean reversion cannot explain the disposition effect either. In their experimental study, subjects forced to sell the losers and given a chance to buy them back usually refuse to do so.

[21] Prospect theory generates individual trading behavior that in equilibrium can explain various asset pricing puzzles. For instance, Benartzi and Thaler (1995) use prospect theory to explain the equity premium puzzle. Barberis, Huang and Santos (2001) show that prospect theory preferences, combined with changes in risk attitudes after prior gains and losses can generate high average stock return, high volatility, and cross-sectional predictability. Grinblatt and Han (2005) suggest that the undervaluation of stocks after gains and overvaluation of stocks after losses generated by prospect theory can predict short-run momentum in returns if the distorted prices are corrected by rational investors. In a study on the trading behavior of Chicago Board of Trade proprietary traders, Coval and Tyler (2005) confirm Grinblatt and Han's (2005) predictions. These studies all assume, implicitly or explicitly, that reference

The literature on prospect theory equates the reference point to the status quo (e.g. Shefrin and Statman (1985), Kahneman, Knetsch and Thaler (1990), Benartzi and Thaler (1995), and Genesove and Mayer (2001)). In this setting the status quo reduces to the average purchase price or equivalently, measuring the reference point in terms of gains from investing in a given stock as I shall do here, zero gains. [22] [23] The literature on the disposition effect often ignores nonlinear probability weighting for simplicity. The informal explanations have so far focused on diminishing sensitivity, which implies that investors in the gains domain are more risk averse, hence more likely to sell a stock (for less risky cash); while investors in the losses domain are risk seeking, hence more willing to hold a stock. However, loss aversion also influences attitudes toward risk in a way that has the potential to explain the disposition effect. When an investor's probability of crossing the reference point is nonnegligible, the kink associated with loss aversion causes first-order risk aversion (Rabin (2000)), potentially decreasing the investor' probability of holding a stock much more than any plausible effect of diminishing sensitivity.[24]

Translating an explanation of the disposition effect based on prospect theory into a formal model has been challenging (e.g. Hens and Vlcek (2005), Gomes (2005),

---

point is determined by the status quo.

[22] In terms of the literature on the disposition effect, Health, Huddart and Lang (1999) is an exception. Because they look at the stock option exercise so there is no natural purchase price to rely on. They show that reaching the highest price of the previous year drives the exercise decision.

[23] A reference point defined by the status quo equates the monetary gains (capital gains and dividends) from investing in a stock to the psychological gains relative to the reference point. However, such equivalence breaks down when the reference point is different from the status quo, in which case I shall use "psychological gains" to describe the latter. Further, reference point is defined on wealth space in prospect theory, but it can be more conveniently referred to in terms of the corresponding reference level of gains in the stock market setting.

[24] Decision makers exhibiting first-order risk aversion are risk averse even for very small gambles; while those with second-order risk aversion, represented by the usual concave utility function, are approximately risk neutral for small gambles.

and Barberis and Xiong (2009)). Barberis and Xiong (henceforth "BX") (2009) propose a dynamic model of selling behavior based on both loss aversion and diminishing sensitivity, taking into account that investors' rational expected returns must be positive to justify the initial purchase decision. BX show that for a binomial or lognormal returns distribution with a reasonable range of positive means, taking the status quo as the reference point, prospect theory in Tversky and Kahneman (1992)'s parameterization actually generates the opposite of the disposition effect in most cases.[25]

While diminishing sensitivity may contribute to the disposition effect in BX's model, the result that generates the opposite of the disposition effect comes from loss aversion. The first-order risk aversion caused by loss aversion suggests that the closer gains are to the reference level, the more risk-averse investors are, hence the more likely they are to sell. Returns distribution with positive mean normally generates large gains and small losses, making gains on average farther away from zero than losses. If an investor's reference level of gains is zero, then she is more likely to sell stocks that currently show a loss because of the proximity to the zero reference level. Loss aversion, even though partially offset by the effect of diminishing sensitivity, therefore generates more sales below than above zero gains.

Although BX carefully investigate the robustness of their results in several directions, they do not consider alternative specifications of the reference point

[25] BX's original formulation takes the initial wealth invested in a given stock (with interest earnings) as the reference point. In section III of the paper they also sketch a model of realization utility that distinguishes between paper and realized gains and losses. Realization utility is capable of generating the disposition effect. I will discuss this alternative in section IV in light of the empirical facts documented in this paper. My analysis, however, still follows the traditional assumption *not* to distinguish between paper and realized gains and losses.

beyond the status quo. However, the empirical literature on prospect theory has taken equivocal positions on what determines reference point. Although the early literature assumes that reference point is the status quo, Kahneman and Tversky (1979) also note that "there are situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo". Kőszegi and Rabin (2006, 2007, 2009) develop this idea in their new reference-dependent model by endogenizing reference points as lagged rational expectations.[26]

This paper reconsiders the possibility of explaining the disposition effect via loss aversion, taking a broader view of the reference point. When the reference point is not closely tied to the status quo, positive expected return and the associated asymmetry of gains and losses around zero play a less important role for the disposition effect, and I show that for a general returns distribution with positive expected return, loss aversion with reference point defined by expectations reliably implies a disposition effect of the kind commonly observed. Diminishing sensitivity reinforces this effect, but it is not essential for an explanation.

Econometric analysis of Odean's (1999) data on individual trading records from a large brokerage house confirms the existence of a large disposition effect. More

[26] Several empirical papers have tested Kőszegi and Rabin's assumption that reference point is determined by expectations. In a lab experiment, Abeler, Falk, Goette, and Huffman (forthcoming) manipulate subjects' expectations of earnings and show that their labor supply decision is determined by reference point defined by expectations rather than the status quo. Ericson and Fuster (2009) suggest that reference point defined by expectations plays an important role in driving the classical endowment effect. Gill and Prowse (2009) run a field experiment in a real effort competition setting. Their subjects respond negatively to the rivals' efforts, a prediction from disappointment aversion that treats the certainty equivalent of the lottery—a plausible proxy of expectations—as the reference point. Crawford and Meng (2009) proxy expectations by natural sample averages of the outcomes to show that a reference-dependent model with reference point defined by expectations provides a useful account of New York cabdrivers' labor supply behavior. Card and Dahl (2009)'s empirical analysis suggests that emotional cues generated by unexpected losses by the home team in football increase family violence.

importantly, a novel and stronger empirical pattern is documented and linked to loss aversion. Being the first attempt to estimate investors' reference points from individual trading data, the econometric analysis also supports expectations as the most reasonable candidate for investors' reference points.

The rest of the paper is organized as follows.

Section 2.2 proposes a model of reference-dependent preferences. I model investors' decision problem much as BX (2009) do, but with the following differences: First, I assume loss aversion but not diminishing sensitivity. Since diminishing sensitivity has explanatory power for the disposition effect, ignoring it strengthens my main point, that prospect theory can provide a credible explanation for the disposition effect. Second, instead of equating the reference point to the status quo, I derive the model's implications for any exogenous and deterministic reference point. Third, I generalize returns distribution from BX's binomial or lognormal to any continuous distribution, so that the model can deliver a complete picture of the non-monotonic changes in risk attitudes across all return levels. Finally, although BX's model is dynamic, to keep the matter simple I illustrate the main prediction of loss aversion using a static model.[27]

The major prediction of the model is a V-shaped relationship between the optimal position on a given stock and current gains of that stock, with the bottom point of the V shape closely linked to the reference point. When we change an investor's reference point from the status quo to expected gain, which should be positive due to the initial purchase decision, the bottom of the V shape changes from zero to a positive

[27] Section 2.2.3 analyzes a dynamic model and a model of stochastic reference point as robustness check.

level. Although most gains are still farther away from zero than most losses, most gains are now generally closer to the positive reference level. Thus the investor is more likely to sell when stock price appreciates from the average purchase price. The V-shaped relationship, combined with the effect of positive expectations on the reference point, therefore generates the disposition effect. However, the V shape is a stronger testable prediction. For instance, a threshold strategy of selling the stocks once certain positive threshold is reached is also capable of generating the disposition effect, but will not yield a V-shaped relationship.

Section 2.3 analyzes individual trading records from a large brokerage house used by Odean (1999) and Barber and Odean (2000, 2001). I pool observations across investors to estimate the aggregate trading pattern. Since investors in the dataset hold either all or none of their positions on a given stock most of the times, the probability of holding a stock becomes a good proxy for the normalized stock position.[28] I indeed document a novel and largely V-shaped relationship between the probability of holding a stock and current gains of that stock. Several papers have partially characterized the implications of this relationship using various datasets, but to my knowledge this paper is the first to document and analyze the complete quantitative pattern.[29]

---

[28] Each investor's position on a given stock in the portfolio at different times need to be normalized relative to the initial position to facilitate the analysis across stocks and investors.

[29] Using a dataset of individual trading records different from the one used in this paper, Odean (1998) shows that investors are more likely to hold big winners and losers than small ones. Grinblatt and Keloharju (2001) find the same tendency with losers using trading records of Finnish investors. Working with the same data set as this paper, Ivković, Poterba and Weisbenner (2005) find that the relationship between the probability of holding a stock and positive capital gains is negative within six months and positive after twelve months since purchase, a natural implication of the V-shaped relationship with positive expectations as the reference point, because in the domain of positive gains, small (large) gains are located to the left (right) of the reference level, leading to a negative (positive)

My theoretical characterization of the V-shaped relationship also suggests a useful empirical strategy to identify the reference point from the bottom of the V. Using a multi-threshold model and treating investors as if they had homogenous reference level of gains, I estimate the bottom of the V shape to be around a gain of 5.5%, which is significantly different from zero, suggesting that the status quo cannot be a reasonable candidate of the reference point. The estimate is, however, closely tied to investors' expectations. Qualitatively, investors' expected gains should be strictly positive to justify their initial purchase decision. Quantitatively, investors' expected gains should be reasonably related to market returns. For the average holding period (230 days) in the sample, the market return is 4.8% five years prior to the sample period and 5.9% during the sample period, both very close to the 5.5% estimate.

Section 2.4 extends the empirical analysis to allow heterogeneous reference points, particularly between (i) frequent traders and infrequent traders, and (ii) stocks with good price history and bad price history. A reference point determined by the status quo does not predict systematic differences along these dimensions but the one determined by expectations does: First, compared to infrequent traders, frequent traders should have a lower bottom of the V, and their relationship between the probability of holding a stock and gains from the stock should be closer to a perfect V shape. These predictions follow because frequent traders hold their stocks for only a short period, which makes their expected gains lower (given positive expected returns) and closely bunching together. Second, controlling for changes in beliefs about future stock returns, stocks with mostly appreciating prices after purchase should

relationship.

demonstrate a higher bottom of the V, simply because past returns are part of the total gains expected from investing in a stock, even in the absence of learning. These predictions are confirmed by the data.

Section 2.5 concludes the paper.

2.2. Theoretical Model

2.2.1. Set-up

The section looks at a static wealth allocation problem in which an investor has to decide on how to split her wealth between a risky asset (stock) and cash. For simplicity I assume no short selling, no time discounting, and no return and inflation risk on cash. The net return of the stock is r_t, $t = 1,2$, which is independently and identically distributed with a continuous distribution $f(r_t)$ on the support $(-1,+\infty)$. The investor starts with a given initial wealth W_0 out of which $x_0 P_0$ is allocated to the stock, and the rest to cash, where x_0 is the number of shares and P_0 is the purchase price. $P_t = P_{t-1}(1+r_t)$ and $g_t = (P_t - \overline{P}_{t-1})/\overline{P}_{t-1}$ denote price and gain from the stock in period t respectively, where $\overline{P}_{t-1}$ is the average purchase price at the end of period $t-1$. To keep the model static, I treat the initial position as given here, but as Barberis and Xiong (2009) and Hens and Vlcek (2005) suggest, I impose positive expected return $E(r_t) > 0$ to reflect the restriction on beliefs about returns implied by the initial purchase decision. Section 2.2.3 discusses the dynamic problem and the implications of including the initial decision into the analysis.

In period one, $P_1 = P_0(1+r_1)$ is realized. The investor chooses a stock position x_1 to maximize the gain-loss utility in period two given the reference point W_1^{RP} and subject to the budget constraint $0 \le x_1 P_1 \le W_0 + x_0 P_0 r_1$. In period one, $g_1 = r_1$, because the average purchase price at the end of period zero is simply the initial purchase price $\overline{P}_0 = P_0$.

In period two, $P_2 = P_1(1+r_2)$ is realized. The investor incurs the gain-loss utility over changes in wealth relative to the reference point W_1^{RP}. Equation (2.1) specifies the expected gain-loss utility in period two. [30] [31]

(2.1) $$E(U(W_2 \mid W_1^{RP})) = E(1_{\{W_2 - W_1^{RP} > 0\}}(W_2 - W_1^{RP}) + \lambda 1_{\{W_2 - W_1^{RP} \le 0\}}(W_2 - W_1^{RP}))$$

(2.2) $$W_2 = W_1 + x_1 P_1 r_2 = W_1 + x_1 P_0 (1+g_1) r_2$$

(2.3) $$W_1^{RP} = W_0 + x_0 \overline{P}_0 g_1^{RP} = W_0 + x_0 P_0 g_1^{RP}$$

W_2 is the wealth level in period two. W_1^{RP} is the deterministic reference point relevant for period-one decision and period-two utility. The reference point is lagged in the sense that it adjusts to neither the price realization P_1 nor the position x_1 in

---

[30] Following BX (2009), this paper defines utility directly over wealth. Equation (2.1) can be understood as an implicit function that reflects utility from an optimal consumption plan in the future given certain wealth level today. The fact that wealth fluctuations compared to a reference point generate utility today can be motivated by the implied changes in future consumption plan relative to a reference point, a concept that Kőszegi and Rabin (2009) term as "prospective gain-loss utility".

[31] Kőszegi and Rabin (2006) develop a more general version of the reference-dependent model in which the total utility is a weighted average of the consumption utility and gain-loss utility, and the unit for gain-loss comparison is also the consumption utility:

$$E(U(W_2 \mid W_1^{RP})) = E(u(W_2) + \eta(1_{\{W_2 - W_1^{RP} > 0\}}(u(W_2) - u(W_1^{RP})) + \lambda 1_{\{W_2 - W_1^{RP} \le 0\}}(u(W_2) - u(W_1^{RP}))))\,.$$

$u(.)$ is the traditional consumption utility and η is the weight attached to the gain-loss utility. This more general version keeps the essentials of loss aversion while incorporating the effect of standard consumption utility. Equation (2.1) can be viewed as a simplified version in which the consumption utility is linear and it has a negligible weight. Having a concave consumption utility function does not change the V shape, but it shifts the bottom of the V shape away from the reference point. Having a non-zero weight on the consumption utility does not change the V shape either, since it only affects the gain-loss utility quantitatively but not the qualitatively.

period one, reflecting the possibility that the investor cannot make peace with the current situation immediately. It is perhaps realistic and certainly convenient to focus on g_1^{RP}, the reference level of gains in period two from period one's perspective, also the level of gains in period one at which the investor can simply sell the entire stock position and reach the reference point in period two. $g_1^{RP} = 0$ corresponds to the status quo assumption while $g_1^{RP} = g_1^E = E((1+r_1)(1+r_2)) - 1$ treats the lagged expected gain as the reference level. [32] Positive expected return $E(r_t) > 0$ ensures positive expected gain $E(g_1^E) > 0$. The loss indicator $1_{\{W_2 - W_1^{RP} \le 0\}}$ takes the value "one" if there is a loss relative to the reference point ($W_2 - W_1^{RP} \le 0$), otherwise "zero". If wealth in period two falls below the reference point, their difference is multiplied by $\lambda > 1$, representing loss aversion. Without assuming diminishing sensitivity, the utility function is piece-wise linear. Thus the kink at the reference point characteristic of loss aversion is the only source of risk aversion.

Following BX, my model makes a non-trivial assumption called "narrow framing" or "mental accounting" (Thaler 1990). First of all, I assume that the investor opens a mental account for each stock after purchase and closes the account once the stock is sold. Thus she incurs gain-loss utility at the individual stock level rather than the portfolio level. Correspondingly, W_0 can be viewed as the maximum amount of

[32] This paper's specification of expectations as the reference point departs from Kőszegi and Rabin's (2006) model in the following aspects. First of all, reference point here is specified as the mean expectation rather than the whole stochastic distribution. It turns out that depending on the returns distribution, stochastic reference point may or may not affect the quasiconvex relationship between the optimal position in a stock and gains from that stock (see section 2.2 and Appendix 2.3). Second, the reference point is exogenous to g_1 and x_1. Endogenous reference point as in Kőszegi and Rabin's (2006) "personal equilibrium" does not generally lead to the V shape observed in the data. For the purposes of explaining the disposition effect and distinguishing between the status quo and expectations, it is therefore a better choice to start with a deterministic and exogenous reference point.

money that she is willing to lose on a particular stock (Barberis, Huang and Thaler (2006), and Barberis and Xiong (2009)). Barberis and Huang (2001) show that treating trading decision as if investors were considering each stock separately fits the empirical facts better than including portfolio choice. Odean (1998) also shows that the disposition effect cannot be explained by portfolio concern. Second, I assume that the investor evaluates her investment outcomes and incurs gain-loss utility over a certain narrow period. Benartzi and Thaler (1995) call it "myopic loss aversion". According to their estimation from the aggregate stock returns, the average evaluation period is about one year. My static model can thus be viewed as describing the optimal decision within one such evaluation period.

I do not take a position on what the reference point is when solving the model. Instead I derive the model's predictions for any exogenous and deterministic reference point, in preparation for section 2.3's econometric analysis, where the model will be used to infer its location from the patterns in the data.

2.2.2. Solution

Loss aversion introduces a cut-off point that divides future return r_2 into those generating gains and those generating losses relative to the reference level, which are assigned weight 1 and $\lambda > 1$ respectively. The cut-off point $K(x_1) = \frac{x_0}{x_1}(\frac{1+g_1^{RP}}{1+g_1} - 1)$ is a function of current position x_1, given the reference level g_1^{RP} and current gain g_1. $K(x_1)$ is a specific level of r_2 that makes period two wealth equal to the reference point $(W_2 = W_1^{RP})$. Equation (2.4) gives the expected marginal utility of holding an additional share. [33]

$$E(MU(K(x_1))) = P_1\left(\int_{K(x_1)}^{+\infty} r_2 f(r_2)dr_2 + \int_{-1}^{K(x_1)} \lambda r_2 f(r_2)dr_2\right) \tag{2.4}$$

The optimal interior solution x_1^* satisfies the first-order condition $E(MU(K(x_1^*)))=0$. Under the piece-wise linear assumption, the choice variable x_1 enters the first-order condition only through $K(x_1)$, which is the sufficient statistic for the optimal solution. Proposition 1 gives the optimal interior position. The restrictions imposed on the returns distribution indicate that returns should be good enough to induce purchase in the first place ($E(r_t)>0$), but they should not be too lucrative to make even a loss-averse investor never want to sell the stock ($E(MU(K(x_1)=0))<0$).

Proposition 2.1. (See Appendix 2.1 for proof) For any returns distribution $f(r_t)$ satisfying $E(r_t)>0$ and $E(MU(K(x_1)=0))<0$, there exist two deterministic return levels $K_1<0$ and $K_2>0$ that satisfy $E(MU(K_1))=E(MU(K_2))=0$. The optimal interior position is given by

$$x_1^* = \begin{cases} \frac{x_0}{K_2}\left(\frac{1+g_1^{RP}}{1+g_1}-1\right), for\ g_1 < g_1^{RP} \\ \frac{x_0}{K_1}\left(\frac{1+g_1^{RP}}{1+g_1}-1\right), for\ g_1 > g_1^{RP} \\ 0, \qquad for\ g_1 = g_1^{RP} \end{cases} \tag{2.5}$$

The budget constraint $x_1^* P_1 \le (W_0 + x_0 P_0 r_1)$ binds for extremely high and low

[33] Equation (2.4) should also include the effect of the change in x_1 on the cut-off point $K(x_1)$. But this term is zero by the fact that at the return level $K(x_1)$, future wealth equals the reference point $W_2 = W_1^{RP}$. Marginal change in $K(x_1)$ thus brings almost no change to the expected gain-loss utility.

g_1.

With a binomial returns distribution, BX (2009) show that it is optimal for the investor to gamble until the highest wealth leaves her at or near the reference point if she is currently in the losses domain, and vise versa. Proposition 2.1 confirms and generalizes this conclusion with any continuous returns distribution, by showing that the investor will gamble until the wealth generated by a fixed return level $K_2 > 0$ ($K_1 < 0$) reaches the reference point when she is currently in the losses (gains) domain. Given such optimal strategy, it is easy to see how the optimal position changes with different levels of gains.

Corollary 2.1. The optimal interior position x_1^*

(i) decreases in g_1 when $g_1 < g_1^{RP}$ and increases in g_1 when $g_1 > g_1^{RP}$.

(ii) is concave in g_1 when $g_1 < g_1^{RP}$ and convex in g_1 when $g_1 > g_1^{RP}$.

Different levels of current gains generate different distances from the reference level g_1^{RP}. The closer g_1 is g_1^{RP}, the more risk averse the investor is, so she takes few risks by demanding few shares of the stock. Meanwhile she takes more risks by enlarging the position in the stock as g_1 deviates farther from g_1^{RP}. These facts bring a V-shaped relationship between the optimal position and gains from the stock, with the bottom of the V shape reached at the reference level $g_1 = g_1^{RP}$. Such relationship comes from both loss aversion and the monotonic probability of crossing the reference point as the position becomes larger. The optimal position x_1^* is concave when $g_1 < g_1^{RP}$ and convex when $g_1 > g_1^{RP}$ because it is more expensive to purchase more shares when stock price is higher. Consequently, the tendency to enlarge the optimal position is mitigated (exacerbated) as gains become larger (smaller). [34]

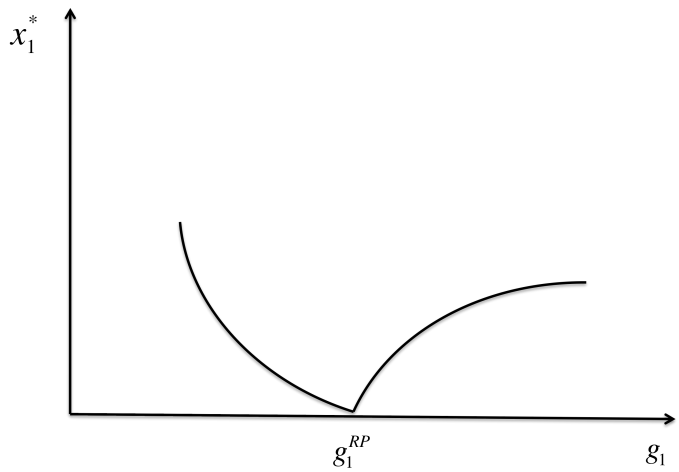


Figure 2.1. The V-Shaped Relationship

This figure shows the relationship between the optimal position x_1^* from a given stock and gains g_1 from that stock for any exogenous and deterministic reference level g_1^{RP}.

Figure 2.1 illustrates the V-shaped relationship in the region where the budget constraint does not bind. The optimal position reaches its minimum with a kink at the reference level $g_1 = g_1^{RP}$. This kink comes from different cut-off return levels K_1 and K_2 used below and above the reference level g_1^{RP}. It is clear that the investor is most likely to sell the stock around g_1^{RP}, given a fixed x_0. For gains located symmetrically around g_1^{RP}, which one leads to a larger optimal position x_1^* is ambiguous.35 However,

[34] Comparative statics analysis shows that higher loss aversion coefficient λ leads to smaller x_1^* at every level of current gains. Loss aversion, like conventional risk aversion, reduces investment demand for the risky asset. Similarly, any move of probability mass from positive returns to negative returns decreases the optimal position at all levels of current gains, but the V-shaped relationship remains unchanged.

[35] The answer depends on where K_1 and K_2 stand relative to zero, which in turn depends on the returns distribution. For example, any returns distribution with an increasing $f(r_t)$ in the region $r_t \in [K_1, K_2]$

since the disposition effect describes asymmetric behavior around zero gains, the asymmetry around the reference level g_1^{RP}, if any, is less relevant in explaining the disposition effect once g_1^{RP} is different from zero.

With the V-shaped relationship, it is convenient to illustrate both BX's argument for why loss aversion fails to predict the disposition effect when zero gains is treated as the reference level (Figure 2.2) and how assuming positive expected gain as the reference level generates the disposition effect (Figure 2.3).

In figure 2.2 BX's argument relies crucially on $g_1^{RP}=0$. Under returns distribution with positive mean, a typical realization of gains (g_1^G) will be relatively farther away from zero compared to that of losses (g_1^L). It therefore takes a larger position for future wealth generated by the cut-off return K_1 to reach the reference point at g_1^G. Diminishing sensitivity in BX's model mitigates this tendency in favor of the disposition effect but is not enough to totally offset it.

As Figure 2.3 illustrates, this paper generates the disposition effect without relying on diminishing sensitivity because of the shift of the reference level of gains from zero to the positive expected gain, making g_1^G closer to g_1^{RP} than g_1^L. Because the investor is most likely to sell around the reference level due to loss aversion, she is on average more likely to sell the stock when it has a gain than when it has a loss.

---

implies $|K_1| > |K_2|$, leading to a relatively small x_1^* at the high gains level in the symmetric pair.

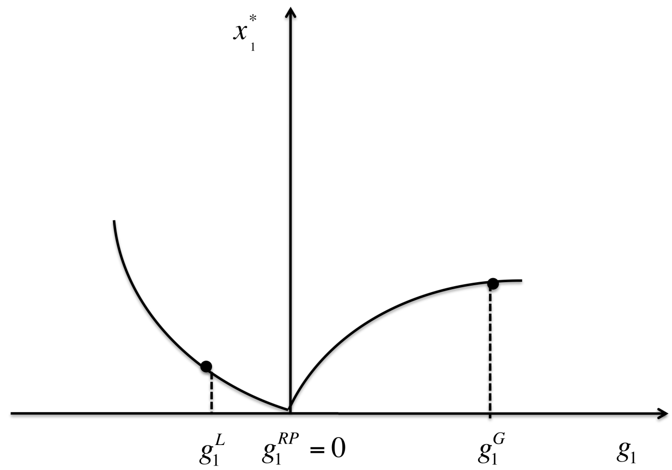


Figure 2.2. The V-Shaped Relationship when the Reference Point is the Status Quo
This figure shows the relationship between the optimal position x_1^* in a given stock and gains g_1 from that stock when the reference level is determined by the status quo ($g_1^{RP} = 0$). g_1^G and g_1^L are the typical realization of gains and losses respectively.

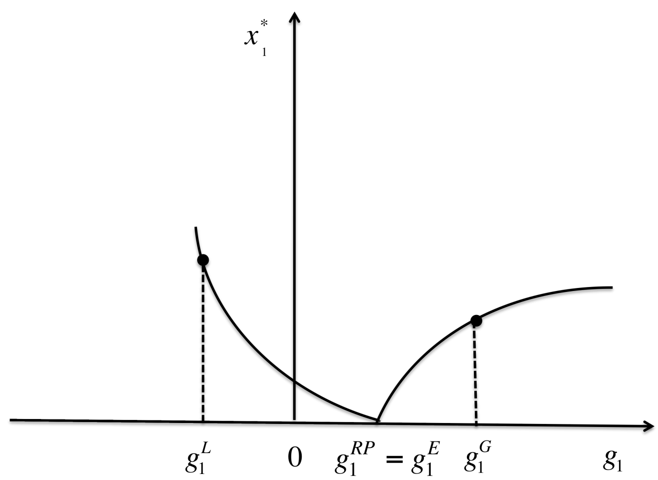


Figure 2.3. The V-Shaped Relationship when the Reference Point is Expectation
This figure shows the relationship between the optimal position x_1^* from a given stock and gains g_1 from that stock when the reference level is determined by expected gain ($g_1^{RP} = g_1^E$). g_1^G and g_1^L are the typical realization of gains and losses respectively.

2.2.3. Robustness

My simple static model predicts a V-shaped relationship between the optimal stock position in a given stock and gains from that stock. This relationship is sufficient but not necessary to generate the disposition effect, because the disposition effect does not require an increasing tendency to hold the stock as gains become large. The model also implies that the bottom of the V shape is the reference level. I discuss the robustness of these results by relaxing one assumption at a time. It turns out that the V-shaped relationship is relatively robust, but not to the endogenous reference point, or the stochastic reference point with certain distributions. Under more general conditions, however, the bottom of the V shape may not be reached exactly at the reference level. But the bottom point is still largely driven by and therefore provides valuable information about the reference point.

- **Multiple periods:** In a reasonable dynamic model the investor has many periods to make choices and incur gain-loss utilities. In the beginning of period t, she learns about P_t (hence g_t) and incurs gain-loss utility relative to the reference point W_{t-1}^{RP}. Then she forms a new (but still lagged) reference point W_t^{RP} for utility in period $t+1$ and chooses x_t accordingly. The investor understands that her choice this period will affect wealth levels and reference points in the following periods. Formally, the investor's decision problem is

(2.6) $$\max_{\{x_0,x_1,\ldots,x_T\}} \sum_{t=0}^{T} E_0(\beta^t U(W_{t+1} \mid W_t^{RP})).$$

(2.7) $$U(W_{t+1} \mid W_t^{RP})) = 1_{\{W_{t+1}-W_t^{RP}>0\}}(W_{t+1}-W_t^{RP}) + \lambda 1_{\{W_{t+1}-W_t^{RP}\le 0\}}(W_{t+1}-W_t^{RP}))$$

(2.8) $$W_{t+1} = W_t + x_t P_t r_{t+1} = W_t + x_t \overline{P}_{t-1}(1+g_t) r_{t+1}$$

(2.9) $$W_t^{RP} = W_{t-1} + x_{t-1}\overline{P}_{t-1} g_t^{RP}$$

The investor chooses the optimal positions $\{x_0^*, x_1^*, ..., x_T^*\}$ to maximize the summation of the gain-loss utilities from period 0 up to a final period $T+1$, given the reference points $\{W_0^{RP}, W_1^{RP}, ..., W_T^{RP}\}$. By definition $x_{T+1} = 0$ and the investor sells the entire position. W_0^{RP} is taken as given before making initial decision, but $\{W_1^{RP}, W_2^{RP}, ..., W_T^{RP}\}$ are partially determined by $\{x_0^*, x_1^*, ..., x_{T-1}^*\}$. I again focus on the sequence of deterministic reference level of gains $\{g_0^{RP}, g_1^{RP}, ..., g_T^{RP}\}$.

Proposition 2.2. (See Appendix 2.2 for proof) In the dynamic model defined above, the optimal interior position x_t^* is decreasing in g_t when $g_t < g_t^{RP}$ and increasing in g_t when $g_t > g_t^{RP}$.

According to proposition 2.2, the general V shape is robust to the dynamic consideration. The bottom point of the V shape is still determined by the reference point. The intuition is the following: Both future wealth levels and reference points adjust to gains g_t in the same direction so their differences (the gains and losses) are not affected by g_t. The current position x_t has a non-zero effect on future gain-loss utilities but such effect is orthogonal to g_t. Therefore the dynamic consideration affects the level of x_t^* but not its relative relationship to g_t.

The dynamic model nonetheless sheds light on the impact of including the initial purchase decision into consideration, which is potentially important because it restricts the range of expected returns we should focus when discussing the subsequent selling decision.[36] As BX correctly suggest, the investor should purchase stocks with only strictly positive expected returns, especially when she is loss averse with respect to the status quo. If her sequence of reference points is determined by positive expectations, however, she is willing to accept stocks with lower (but still positive) expected returns. Because only by taking more risks can she have the opportunity to obtain the desired positive gains. On one hand, since the disposition effect in my model is driven by treating the positive expectations as the reference points, a lower but still positive expected return makes the disposition effect weaker but still present. On the other hand, a lower expected return reduces the asymmetry of gains around zero, which according to BX's argument, favors the disposition effect.

- **Stochastic reference point:** If the investor's reference point is stochastic in nature, the overall utility is a probability-weighted average of the gain-loss utilities relative to different realization of the reference point.

Proposition 2.3. (See Appendix 2.3 for proof) For a stochastic reference level of gains g_1^{RP} with the density $h(g_1^{RP})$, there exists a lower bound $\underline{g}_1$ and a higher bound

[36] For a consistent investor who believes that the returns distribution is independently and identically distributed, if the reference level of gains is zero and remains so over time, once she purchases the stock she will never want to sell. This theoretical reasoning makes it unlikely for the status quo to be the reference point in my simple static model. The problem of never wanting to sell does not exist under the additional assumptions such as dynamic adjustment of the reference point, diminishing sensitivity and stochastic reference points. My empirical analysis, however, does not impose the theoretical structure on estimation, so the econometric model is free to pick up zero as the reference level, if the data suggests so. If I indeed estimate a bottom point of the V shape at zero, then I need to modify my static model to address the problem of never wanting to sell. However, given that this is not happening in the data, this issue seems to be minor.

$\overline{g}_1$ such that: the optimal position x_1^* is decreasing in g_1 for $g_1 < \underline{g}_1$ and increasing in g_1 for $g_1 > \overline{g}_1$. The relationship is ambiguous when $\underline{g}_1 \le g_1 \le \overline{g}_1$.

Although the V shape does not literally exist given stochastic reference point, its essential nature is still preserved. However, there is hardly any bottom point, global or local, that can be identified with the reference point. A highly relevant case under the stochastic reference point is the possibility of comparing outcomes to both the status quo and expectations. Depending on the returns distribution $f(r_2)$ and the probability attached to each reference point, there are two possible patterns. Under the "single-trough" pattern the optimal position is still V-shaped but the bottom of the V is located between the status quo and expected gain. Under the "twin-trough" pattern the optimal position is W-shaped, with a local minimum position at each reference level.

- **Endogenous reference point:** An endogenous reference point that is affected by the decision variable x_1 does *not* create the kind of variations in psychological gains and losses as x_1 and P_1 vary that is necessary to keep the V shape. [37] It is interesting to test the predictions of endogenous reference point, but given a strong V-shaped pattern in the data, it seems more natural to explore the possibilities of models that explain this pattern with exogenous reference point first.

- **Diminishing sensitivity:** Making the gain-loss utility function concavity above and convexity below the reference point as Tversky and Kahneman (1992) estimate them keeps the V-shaped relationship, because diminishing sensitivity itself generates monotonically decreasing position in a given stock as gains from that stock

[37] Gill and Prowse (2009) provide the first lab evidence for the existence of endogenous reference point

increase. The bottom of the V shape, however, is not exactly at the reference level. Instead it is pushed slightly to the right of its reference level (see Gome (2005), and Barberis and Xiong (2009)). Since the gain-loss utility function with diminishing sensitivity is not globally concave, there is a discontinuous decline in the optimal position at the bottom of the V shape (Gomes (2005)), partially reflecting the power of diminishing sensitivity to explain the disposition effect.

2.3. Empirical Analysis: Overall Sample

2.3.1. Data Summary

The dataset is from a large discount brokerage house on the investments of 78,000 households from January 1991 through December 1996. It was used by Odean (1999) and Barber and Odean (2000,2001) but is different from the one that Odean (1998) used in his pioneering analysis of the disposition effect. The original dataset includes end-of-month position statements, trading records (trade date, trade quantity and trade price) for each stock held by each account, together with some background information about the account owners (e.g. gender, age, income and net wealth). Odean (1999) gives more detailed description of this dataset. I also have data for the daily stock price, daily trading volume, shares outstanding, an adjustment factor for splits and dividends and market returns (S&P) from CRSP. Appendix 2.4 reports my data cleaning process.

In this study I focus on the trading of common stocks at the individual stock level, ignoring portfolio concern. My empirical analysis relies on constructing an investor's trading history for each stock she holds. The history includes dates of purchase, hold and sale. Following Odean (1998) and Grinblatt and Keloharju (2001), I generate observations of hold dates in the following way: Any time that at least one sale takes place in the portfolio, I count the untraded stocks in the portfolio as holds and obtain price information from CRSP for them. These are stocks that investors could have sold but did not. In other words, I select holding dates conditional on having at least one sale taking place in the portfolio on that day. This procedure is standard in the empirical literature of the disposition effect. It ensures that any holding decision in the sample comes from deliberation rather than inattention.[38]

Table 2.1 reports the summary statistics of the dataset. All prices are appropriately adjusted for commission, dividend and splits so the calculated gains from a stock include both capital gains and other forms of income.[39]

Table 2.1. Summary Statistics

	Mean	Standard Deviations	Min	Max	Observations
Holding Days	230	290	1	2126	279,968
Paper Gains	0.005	0.353	-1	24.01	848,756
Realized Gains	0.051	0.371	-1	21.55	279,968
Portfolio Size	4.044	5.677	1	309	1,128,724
Principal at the Initial Purchase	10,622	28,302	6	6,011,361	394,637
Commission per Share	0.448	1.319	0	55	394,637
Income	95,851	3,728,442	500	588,671,000	36,174
Age	55	14	1	80	18,724

For each investor, I construct the variable "trading frequency" as the inverse of

[38] For the purpose of studying the relationship between the optimal position in a stock and gains from that stock, it does not necessarily imply sample selection problem. However, there is some risk of bias if portfolio size and trading frequency are correlated to gains. I control for these potential confounding factors in the regression.

[39] Whether to adjust for commissions doesn't generate a big difference in the estimates.

the average day between two trades using her entire trading records. Investors with trading frequency higher (lower) than the mean are treated as frequent (infrequent) traders. Frequent traders on average trade every month and infrequent traders trade every five months. To measure the desirability of price history during the holding period, I calculate the proportion of increasing prices (compared to the price of the previous observation) between the initial purchase date and the date in question, and assign observations with this ratio higher (lower) than the mean as having good (bad) history, in the sense that the stock price keeps going up (down) on average after the initial purchase.[40] Stocks with good price history yield an average gain of 22.4% and those with bad price history yield an average loss of 16.3%. For both sample-splitting criterions, ties are randomly assigned to each group.

I follow Odean's (1998) method of measuring the disposition effect to analyze his (1999) dataset and report the results in Table 2.2. Odean (1998) count winners and losers relative to the average purchase price.[41] He defines the proportion of gains realized (hence PGR) as the number of realized gains divided by the number of realized and paper gains. Similarly the proportion of losses realized (hence PLR) is defined as the number of realized losses divided by the number of realized and paper losses.

[40] I have tried other ways to measure the desirability of price history, including constructing the proportion of positive gains and the proportion of positive market adjusted gains between the initial purchase date and the date in question. To check whether investors judge good or bad history by more recent history, I also calculate these ratios for the past week or past two months. The qualitative results of this paper are robust to these alternative specifications.

[41] A hold observation is counted as a gain (loss) if the lowest (highest) price of that day is higher (lower) than the average purchase price.

Table 2.2. Proportion of Losses Realized (PLR) and Proportion of Gains Realized (PGR)

	Overall Sample	Dec.	Jan.-Nov.	Frequent Traders	Infrequen t Traders	Good History	Bad History
PGR	0.360	0.309	0.365	0.227	0.424	0.380	0.218
PLR	0.245	0.313	0.238	0.152	0.304	0.248	0.244
PLR-PGR	-0.115	0.004	-0.127	-0.076	-0.120	-0.132	0.026
t-statistic	-130.054	1.326	-137.925	-58.724	-104.358	-69.649	14.496
RG	184,802	11,872	172,930	37,475	147,327	171,456	13,346
RG+PG	512,647	38, 434	474, 213	164,974	347,673	451,519	61,128
RL	136,041	15,572	120,469	32,780	103,261	15,051	120,990
RL+PL	556,375	49,741	506,634	216, 325	340,050	60,746	495,629

Note:
This table calculates the proportion of gains realized (PGR) and the proportion of losses realized (PLR) following the strategy of Odean (1998 Table I). PGR is calculated as the ratio between numbers of realized gains and total (realized and paper) gains; PLR is calculated as ratio between numbers of realized losses and total (realized and paper) losses. RG, PG, RL, PL represent numbers of realized gains, paper gains, realized losses and paper losses. The standard error for the t-statistic is constructed by $\sqrt{\frac{PGR(1-PGR)}{RG+PG}+\frac{PLR(1-PLR)}{RL+PL}}$.

My dataset does exhibit substantial disposition effect. In the overall sample, investors realize significantly more gains than losses. The difference is large (PLR-PGR=-0.115). [42] In December the difference is not significant, probably due to tax incentive to realize more losses than gains.[43] The difference between frequent traders and infrequent traders is also consistent with Odean's finding. Although both types demonstrate the disposition effect, infrequent traders are especially vulnerable to it.[44]

[42] The difference between PLR and PGR for the overall sample is -0.05 in Odean (1998 Table I), smaller than the one reported in Table 2.2. The qualitative nature is nonetheless consistent across the two samples.

[43] Because of tax on positive capital gains, it is often more beneficial for investors to realize losses than gains at the end of the year. Ivković, Poterba and Weisbenner (2005) identify the existence of tax-driven selling using the dataset in this paper.

[44] BX (2009) explain this fact by noting that infrequent traders have long average holding period, so future independent risks can cancel each other and make infrequent traders more willing to accept stocks with lower expected returns today. Gains and losses are then distributed more symmetrically

The distinction between good history and bad history, which has not been investigated before, also reveals a surprising asymmetry. Investors with good price history are more likely to sell stocks with gains and hold on to those with losses. The pattern is reversed in the case of bad history.

2.3.2. The Probability of Holding a Stock: Illustration

While my model predicts a relationship at the individual level, this section's empirical analysis pools the observations across investors. For comparison purpose, investors' positions on stocks are normalized relative to the initial purchasing positions. In the individual trading records constructed from my dataset, only 4% of the observations are partial sales or repurchases. In other words, investors sell either all or none of their positions most of the times. This fact makes the normalized stock positions essentially binary, and so I can further proxy the normalized optimal position of account i, stock j and time t using a binary "hold" variable h_{ijt} ($h_{ijt}=0$ if sell and $h_{ijt}=1$ otherwise).[45] Also let $g_{ijt}=P_{ijt}/\overline{P}_{ijt}-1$ denote account i's gain from investing in stock j at time t, where $\overline{P}_{ijt}$ is the average purchase price. Because of the almost binary nature of the decision in the sample, $\overline{P}_{ijt}$ takes constant values at the initial purchase price most of the times.

Figure 2.4 illustrates the probability of holding a stock calculated as the average of the dummy variable h_{ijt} within each 10% gains interval. The patterns in

around zero, making their model more likely to generate the disposition effect. I develop an alternative explanation in section 2.4 based on the effect of the reference points defined by expectations.

[45] For comparison purpose h_{ijt} is purposefully constructed to be exactly one minus the "sell" dummy variable normally used in the literature.

this and the upcoming figures are shown to be robust to the influences of trading frequency, portfolio size, and the stock's own returns and market returns in the past two months, etc. (see the full set of control variables in Table 2.3.A in Appendix 2.5)

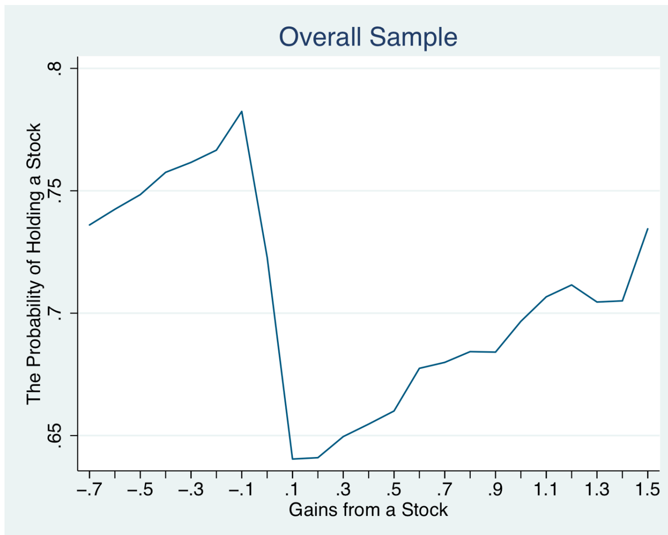


Figure 2.4. The Probability of Holding a Stock (Overall Sample)

This figure shows the probability of holding a stock across different levels of gains for the overall sample. The probability of holding a stock is calculated as the average of a dummy variable h_{ijt} ($h_{ijt} = 0$ if sell and $h_{ijt} = 1$ otherwise)) within each 10% gains interval.

The relationship is clearly non-monotonic. Starting from a loss of 10% it is also V-shaped: The probability of holding a stock starts to decline from a loss of 10% to a gain of 10% and rises after a gain of 10%. The overall pattern implies that investors are most likely to sell stocks with small positive gains and hold stocks with large gains and losses. Such pattern implies the disposition effect, but it is a stronger empirical regularity. The left-tail drop in the probability of holding a stock is not directly predicted by loss aversion, and I will discuss possible explanations in section 2.4.

Figure 2.4 suggests that investors' risk attitudes change non-monotonically as gains increase. Such sharp and non-monotonic change in risk attitudes is not readily explained by expected utility model. The commonly used utility functions (e.g. CRRA) normally imply a monotonic change in risk attitudes as wealth increases, hence a monotonic relationship in Figure 2.4. [46] This largely V-shaped pattern is nonetheless consistent with a model of loss aversion. Further, it is visually clear that the bottom of the V shape is reached at a strictly positive level of gains. This is strong evidence that investors' reference points are affected by something higher than the status quo.[47]

Figure 2.4(S) is a "zoom-in" version of Figure 2.4 on the interval [-10%, 10%], with h_{ijt} averaged for each 0.1% gains interval to calculate the probability of holding a stock. Interestingly, although the bottom of the V shape is not reached at zero gains, there is a steep decline in the probability of holding there. This observation suggests the role of the status quo is not completely negligible. My econometric model includes a dummy variable for positive gains to capture the effect of the status quo.[48]

---

[46] One may argue that because of diversification of risks, gains from one stock in the portfolio may be offset by losses from other stocks so that the effect of gains from individual stock on total wealth is ambiguous, making Figure 2.4 an inaccurate reflection of change in risk attitudes across different wealth levels. This point is well taken. But it should be noted that to rely on this particular point to explain the non-monotonic relationship observed, the expected utility model requires a complicated and specific pattern of correlation among prices of stocks in the portfolios, which may be true but needs more careful study.

[47] In two cases a reference point determined by the status quo can generate a bottom of the V shape located at a positive level of gains. The first one is loss aversion combined with diminishing sensitivity (Gomes 2005), but it cannot be an empirically plausible explanation because it fails to predict the disposition effect (Barberis and Xiong 2009). The second one is Kőszegi and Rabin's (2006) more general reference-dependent model with a concave consumption utility to serve as the unit of gain-loss comparison. However, the magnitude of such shift is likely to be too small to offset the asymmetric gains and losses realization around zero. So it may fail to generate the disposition effect according to BX's argument. Further, both alternatives have a hard time accounting for heterogeneity in the bottom point of the V shape across trading frequency and price history (section 2.4).

[48] Appendix 2.3 analyzes the possibility of treating both the status quo and expectations as reference points. Depending on the parameters, such model may generate prediction indistinguishable from the model with a single reference point determined by expectations. It is also possible that there are

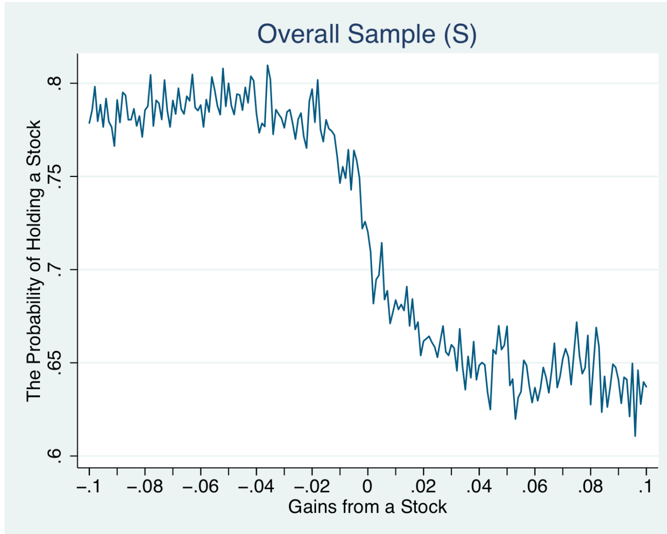


Figure 2.4(S): the Probability of Holding a Stock (Small Region)

This figure shows the probability of holding a stock across different levels of gains for the gains interval [-10%, 10%]. The probability of holding a stock is calculated as the average of a dummy variable h_{ijt} ($h_{ijt} = 0$ if sell and $h_{ijt} = 1$ otherwise)) within each 0.1% gains interval.

2.3.2. The Probability of Holding a Stock: Estimation

This section estimates the non-monotonic relationship shown in Figure 2.4 using a multi-threshold model with unknown thresholds (Andrews (1993), Bai (1996), and Hansen (2000a, 2000b)). A formal estimation is intended to check the robustness of the V shape to the inclusion of more control variables and to obtain accurate estimate of the bottom point of the V shape.

In this section I impose homogeneity in the reference point by treating the location of the bottom of the V shape as a single fixed number. The resulting estimate

heterogeneous investor types, where one type has zero gains while the other type has positive gains as the reference levels. However, in the subsamples generated by trading frequency and price history, reaching zero gains almost always has significant influences over trading behavior. Even if there are different types it seems hard to find an intuitive criterion to sort them out.

reflects the average reference level across investors and stocks in the sample.[49] However, this homogeneity assumption is obviously too restrictive, especially when expectations is a candidate for the reference point. Section 2.4 partially relaxes this assumption by allowing heterogeneity in trading frequency and price history.

The underlying specification is a linear probability model regressing the binary holding decision h_{ijt} on gains from the stocks g_{ijt} and other variables represented by the vector X_{ijt}.[50] I use linear rather than nonlinear specification for the probability mainly because the multi-threshold model I use (Hansen 2000) applies to the linear case. Also significant bias due to probability boundary effects is unlikely because my sample selection of holding observations is conditional on having at least one sale taking place on that day. As the estimation results show, the linear model fits the relationship reasonably well.[51]

Equation (2.10) represents the specification with one threshold. Later I introduce a sequential procedure to estimate multiple thresholds based on this

---

[49] If there is heterogeneity in the threshold, and if it is correlated with the regressors in the equation, then the current specification causes biases to the slope parameters. However, whether it leads to systematic bias on the estimated location of the threshold is unknown. Section 2.4 adds some important heterogeneity to attenuate this problem. To remedy this problem, I have also tried maximum likelihood estimation where I specify the threshold as a linear function of holding period, trading frequency, price history, and a random error. The estimation failed to converge possibly because of the multiple changes in slope. Specifying the threshold as a linear function of these variables without a random component requires developing new tests beyond the scope of Hansen (2000). I leave this to future research.

[50] Another reasonable alternative is to use the Cox proportion hazard model, which has the advantage of controlling for the effect of holding period non-parametrically. However, as far as I know, there is no appropriate procedure to test the location of the threshold in the Cox model. What's more, in my sample holding period has an almost linear effect on the probability of holding a stock, thus a linear model may not cause severe bias. To double check, I include dummy variables for different lengths of holding period to allow nonlinear effect, and the estimation results are essentially unchanged. Although my model implies that the probability of holding is concave below and convex above the bottom of the V shape, I choose to estimate the simple linear model as the first step. More complicated models such as a polynomial, or nonparametric estimation can be used to obtain more accurate estimates.

[51] I include 5th and 95th percentiles of the predicted probability of holding in related tables in Appendix 2.5 to show that most of the predicted values are within the range (0,1).

specification. The reason to allow for multiple thresholds is the observation from Figure 2.4 that besides the slope change at the bottom of the V shape there is also an obvious slope change at a slight negative level of gains. We need to control for such structural change to allow for a better estimate of the bottom point.

(2.10) $$h_{ijt} = \alpha_0 + \alpha_1 I(g_{ijt} > 0) + \beta_0 g_{ijt} + \beta_1 \max(g_{ijt} - g^s, 0) + X_{ijt}\gamma + \varepsilon_{ijt}$$

Equation (2.10) allows the slope of gains to change at an unknown threshold g^s while keeps the probability of holding continuous.[52] I am interested in both the magnitude of the change β_1 and the location of the threshold g^s. $I(g_{ijt} > 0)$ is introduced to capture any influence generated by the distinction between gains and losses. This term controls for the effect of the status quo, and so allows the model to concentrate on finding the reference level via estimation of the bottom point. Further, since my focus is to find the point where slope changes, controlling for the intercept change at zero gains still permits the model to estimate a threshold there, if any. X_{ijt} contains control variables.

A common problem with estimating a threshold model with unknown threshold is the existence of nuisance parameter (see the discussion in Andrews (1993), and Hansen (1996)). Under the null-hypothesis of no change in slope ($\beta_1 = 0$), the threshold g^s does not even exist. The test of any change in parameter value is

[52] Equation (2.10) imposes two restrictions on parameters. First, the slopes of control variables (γ) do not change at the threshold. There are no particular reasons that I can think of for variables such as holding period, portfolio size and past returns to have the threshold effect. So this restriction is made to achieve high efficiency. Second, the regression function is assumed to be continuous at the threshold, because the purpose of the estimation is to find the bottom point of the V shape, which is reflected by a slope change rather than a discrete jump. To check the robustness of the results to these two restrictions, I estimate a model that allows all the coefficients to change at the threshold, including the intercept. The locations of the thresholds are not much different. It also turns out that most control variables do not experience significant changes at the thresholds, except for some variables indicating past returns. There is also no significant discrete jump in the probability of holding at the threshold.

therefore non-standard and normally requires simulation. Card, Mas and Rothstein (forthcoming) propose a simple solution to this problem: They randomly split the data into an estimating sample and a testing sample, using the former to estimate the thresholds and the latter to test the magnitude of parameter changes taking the estimated thresholds as given. This procedure allows for a standard hypothesis testing. I follow their estimation strategy.

I am also interested in testing the hypothesis that $g^s = 0$. Hansen (2000a) constructs a confidence interval for the estimated threshold based on the likelihood ratio test. [53] His test statistic is non-standard but free of nuisance parameter problem. I follow his econometric technique.

Bai (1997) and Bai and Perron (1998) propose a sequential procedure to estimate multiple thresholds with efficiency.[54] The first step is to perform a parameter constancy test to the entire sample and estimate a threshold if the test is rejected. The second step is to split the sample into two subsamples using the estimated threshold from the first step and estimate a threshold in each subsample, if any. Continue this process until no further threshold is detected on each subsample. The third step is to go back and re-estimate the thresholds that are previously estimated using samples containing other thresholds. The third step ensures efficiency. I follow this sequential procedure in estimating multiple thresholds.

Table 2.3 reports the estimates of the key parameters. Column 1 regresses the

[53] He makes an assumption from the change-point literature, which states that as sample size increases to infinity, the change in the parameter value converges to zero. It implies that the statistical test and confidence interval are asymptotically correct if the change in the parameter value is small.

[54] Bai and Perron (1998) construct a model that can estimate and test multiple change points simultaneously. They show that the sequential procedure introduced here is consistent with the simultaneous estimation.

binary holding decision on gains from the stock and a dummy variable that indicates positive gains. Column 2 additionally controls for December effect, holding period, trading frequency, portfolio size, tax rate, income, net wealth, daily trading volume and total shares out. Column 3 further adds market returns and the stock's own returns dating back as far as two months to control for beliefs.[55] β_1 and β_2 are the changes in the slope of gains at threshold I and II respectively. By definition, these changes are all significantly different from zero at 1% significance level. Instead of reporting β_1 and β_2 separately, Table 2.3 reports $\beta_0+\beta_1$ (the slope between threshold I and threshold II) and $\beta_0+\beta_1+\beta_2$ (the slope after threshold II).

The two estimated thresholds are very robust to the inclusion of more control variables, with one estimated at negative levels (-3.4%, -3.9%, -3.9%) and the other at positive levels (5.1%, 5.3%, 5.5%), both significantly different from zero at 1% significance level. Since the estimation results are very similar across columns, I focus on discussing column 3. In the region immediately before threshold I (g_{ijt}<-3.9%) an one unit increase in gains makes investors 15.7% more likely to hold a given stock. After threshold I, the estimated relationship is V-shaped: At threshold II the slope changes from a significantly negative level (-0.812) to a significantly positive level (0.018), making threshold II at 5.5% the bottom point of the V shape. Investors are also 5.5% less likely to hold a given stock once there are positive gains, reflecting the influence of the status quo.

[55] These include the market returns and the stock's own returns in the past 0~1 days, 1~2 days, 2~3 days, 3~4 days, 4~5 days, 6~20 days, 21~40 days, 41~60 days.

Table 2.3. A Multi-threshold Model of the Probability of Holding a Stock (Overall Sample)

	(1)	(2)	(3)
Estimating Sample			
Threshold I	-0.034***	-0.039***	-0.039***
[99% confidence interval]	[-0.043,-0.025]	[-0.047,-0.029]	[-0.048,-0.030]
Threshold II	0.051***	0.053***	0.055***
[99% confidence interval]	[0.034,0.068]	[0.033,0.085]	[0.038,0.078]
Observations	566599	566437	562081
Testing Sample			
β_0	0.084***	0.125***	0.157***
(Slope before threshold I)	(0.005)	(0.005)	(0.010)
$\beta_0 + \beta_1$	-1.149***	-0.960***	-0.812***
(Slope between threshold I and II)	(0.092)	(0.078)	(0.075)
$\beta_0 + \beta_1 + \beta_2$	0.055***	0.013*	0.018***
(Slope after threshold II)	(0.007)	(0.007)	(0.006)
α_1	-0.052***	-0.051***	-0.055***
(Discontinuity at zero)	(0.005)	(0.004)	(0.006)
Market and stock's own returns in the past two months	-	-	Yes
Other control variables	-	Yes	Yes
Adjusted R^2	0.018	0.088	0.094
Observations	562125	561944	557604

Note:
This table reports the estimation results of a multi-threshold model that regresses a binary decision to hold a stock or not on gains from that stock and other control variables. The slope of gains is allowed to change at multiple unknown thresholds. I randomly split observations into an estimating sample (to identify the thresholds) and a testing sample (to test the magnitude of the changes). In the estimating sample I use the procedure developed by Hansen (2000) to construct the heteroskedastic-consistent 99% confidence interval for the location of the threshold based on a likelihood ratio test. In the testing sample, because of the sample splitting the test for the slope change is standard. standard errors clustered by account number are reported in brackets. Estimates of control variables are reported in Table 2.3.A of Appendix 2.5

The estimated bottom point of the V shape (5.5%) is closely related to investors' expectations. Qualitatively, investors' expected gains should be mostly positive to justify the initial purchase decision, and the estimated bottom point of the V shape is significantly different from zero. Quantitatively, for comparison purpose, I calculate the average market returns (S&P) as a proxy for expectations. Over the average holding day in the sample (230 days) the market return is 4.8% in the five years prior to the sample period and 5.9% within the sample period, both very close to an estimated gain of 5.5%. The bottom point of the V shape is also higher than the interest rate. Between the year 1991 and 1996, the return of 1-year Treasury Bill ranges from 3.33% to 5.69%, lower than 5.5% over an average of 230 days.56 Treating expectations as the reference point therefore generates predictions consistent with the empirical estimates from the overall sample.

Most control variables have significant effects as well. The effect of trading frequency is strongly positive, which may sound surprising at first because intuitively frequent traders should be less likely to hold a stock. However, it is a reflection of the sample selection procedure and the trading pattern of frequent traders: Conditional on at least one sale taking place on any given day, frequent traders actually sell only a small proportion of stocks in their portfolios, making the probability of holding a stock on a given day higher. Portfolio size and holding period also have slightly positive and significant effects. Investors are also less likely to hold a given stock in December. High trading volume of the stock on the day makes investors very likely to sell. Income and net wealth do not have very significant effects. Market returns in the past two months mostly affect the probability of holding positively, whereas the stock's own returns in the same period affect the probability of holding negatively.

2.4. Empirical Analysis: Heterogeneity

2.4.1. The Probability of Holding a Stock: Illustration

56 See http://www.federalreserve.gov/releases/h15/data/Annual/discontinued_AH_Y1.txt.

This section investigates heterogeneity by splitting the sample first by frequent traders versus infrequent traders, and then by stocks with good history versus bad history. As the initial step of the analysis, I still keep homogeneity assumption within each subsample by treating the bottom of the V shape as a fixed number, but allow the estimate to vary across subsamples.

To demonstrate the predictions of treating expectations as the reference points, I make the following assumption based on the dynamic model in section 2.2.3:

$$g_t^{RP} = g_t^{E} = \prod_{m=1}^{m=t-1}(1+r_m)E(\prod_{n=t}^{n=T+1}(1+r_n))-1 \tag{2.11}$$

The lagged expected gain g_t^E incorporates the effect of past returns realization up to period $t-1$ and takes expectation over returns from period t to $T+1$. The simple specification is convenient to explain the predicted heterogeneity, but the basic intuition should be robust to a range of more complicated specifications of expectations.

The first heterogeneity is between frequent traders and infrequent traders. I model their difference by different lengths of evaluation period. Frequent traders are assumed to have a relatively short time interval to take an action and evaluate gain-loss utilities.[57] (To reiterate, in the sample frequent traders on average trade every month while infrequent traders every five months). Due to the short evaluation period, frequent traders naturally expect low gains, given $E(r_t) > 0$. If investors' reference

[57] The implicit assumption here is that the decision period and action period have the same length. This assumption can be easily relaxed to the case where investors have multiple periods to adjust their stock positions before evaluating gains and losses. Other things equal, future independent risks in returns would cancel each other, making investors more likely to take larger positions. This channel brings monotonic effect to the levels of the optimal positions for every level of gains. The V shape still remains.

points are affected by expectations, in aggregate the average bottom point of the V shape of frequent traders should be lower than that of infrequent traders.[58]

There is a more subtle prediction. Frequent traders are less likely to adjust their expected gains too much from the initial levels due to the short holding period, while infrequent traders' expected gains may deviate a lot from the initial levels because they experience realization of gains and losses over many periods. The relative bunching of frequent traders' expected gains reinforces the V-shaped relationship, because the V shape is derived from assuming a single reference point. On the contrary, the large variation of infrequent traders' expected gains may not generate a perfect V shape.

The second heterogeneity is between stocks with good history and bad history. Other things equal, investors' expected gains on stocks with rising prices between period 0 and period t should be higher, simply because good price histories generate higher cumulative returns $\prod_{m=1}^{m=t-1}(1+r_m)$, which is part of the expected gains. Admittedly, past returns may change beliefs about the returns distribution $f(r_t)$ if there is learning. Learning either reinforces (the momentum belief) or mitigates (the mean reversion belief) the direct effect of past price history on expected gains. The regression includes the stock's own returns and market returns in the past two months to control for learning.

These predictions are confirmed by the data.

---

[58] It could also be that for some exogenous reasons not related to time horizon some investors expect to earn low gains from holding a stock, and such reference levels make them sell the stock quickly. The prediction of a low bottom point for the sample of frequent traders, however, only requires a correlation (rather than a causality) between time horizon and expected gains.

Figure 2.5 illustrates the probability of holding a stock in the samples of frequent traders as well as infrequent traders. Consistent with the predications, the bottom point of the V shape of frequent traders is indeed lower than that of infrequent traders. Further, frequent traders have a V-shaped relationship that almost perfectly matches the prediction of loss aversion, whereas the pattern of infrequent traders is not strictly V-shaped, mainly because there exists a left-tail drop in the probability of holding a stock.

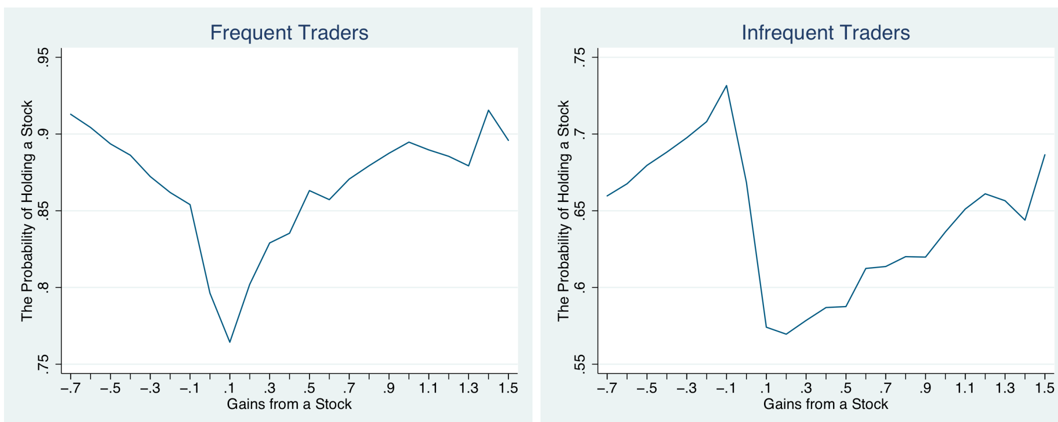
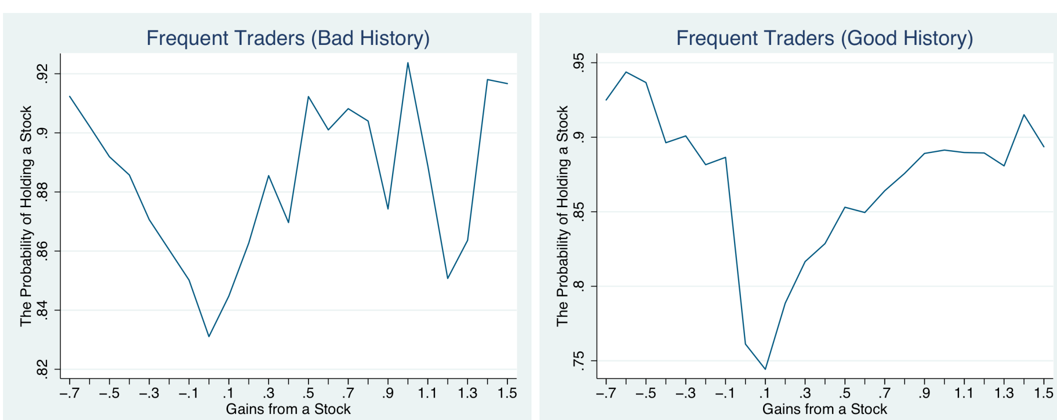


Figure 2.5. The Probability of Holding a Stock (Frequent Traders and Infrequent Traders)

This figure shows the probability of holding a stock across different levels of gains for both frequent and infrequent traders. The probability of holding a stock is calculated as the average of a dummy variable h_{ijt} ($h_{ijt} = 0$ if sell and $h_{ijt} = 1$ otherwise)) within each 10% gains interval.

I further split the observations of each investor type by whether the stocks have good price history or bad price history. Figure 2.6 illustrates the probability of holding a stock in the resulting four subsamples. The two subsamples of frequent traders still keep the basic V shape, except for some noise at tails, because there is a positive correlation between current gains and price history that leads to relatively more observations at the right (left) tail of stocks with good (bad) price history. Interestingly, the bottom point of the V shape in the sample of stocks with bad price history is around zero, while that in the sample of stocks with good price history is positive. The two subsamples of infrequent traders demonstrate large differences. The probability of holding a stock in the sample of stocks with bad price history is almost monotonically increasing as gains increase, with a slight drop when gains become positive. The pattern of stocks with good price history is largely consistent with the V shape prediction, with no obvious left-tail drop in the probability of holding a stock.

It is also informative to compare these patterns across columns. For stocks with bad price history, frequent traders do not increase the tendency to sell as losses become large, but infrequent traders do so. For stocks with good price history, frequent traders have a steep increase in the tendency of holding after reaching the bottom of the V shape while infrequent traders keep the probability of holding almost flat. Recall that the aggregate pattern in Figure 2.4 is largely V-shaped except for the left-tail drop in the probability of holding. It is clear now that such a pattern mainly comes from infrequent traders holding stocks with bad price history.



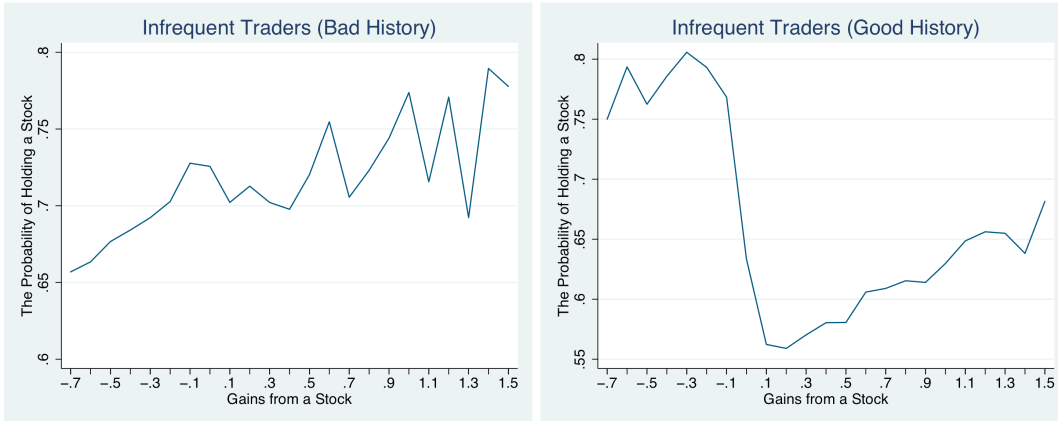


Figure 2.6. The Probability of Holding a Stock
(Stocks with Bad History and Good History)

This figure shows the probability of holding a stock across different levels of gains for stocks with good history and bad history for each trader type. The probability of holding a stock is calculated as the average of a dummy variable h_{ijt} ($h_{ijt} = 0$ if sell and $h_{ijt} = 1$ otherwise)) within each 10% gains interval.

2.4.2. The Probability of Holding a Stock: Estimation

This section estimates the same multi-threshold linear probability model for each subsample. Table 2.4 and Table 2.4 report the estimates of key parameters in the samples of frequent traders and infrequent traders respectively.

Table 2.4 suggests that frequent traders as a whole (column 1) have three thresholds estimated at -2% (threshold I), 1.8% (threshold II) and 26.2% (threshold III) respectively, all significantly different from zero at 1% significance level. The slope of gains changes from a significantly negative level (-1.954) to a significantly positive level (0.234) at threshold II, making it the bottom point of the V shape. Before threshold I and after threshold III the slope is not significantly different from zero. Such pattern is very close to the V shape prediction of the model assuming a fixed reference point. Frequent traders as a whole are also 2.7% less likely to hold a given stock once gains become positive.

Frequent traders holding stocks with bad price history (column 2) have two thresholds estimated at -4.3% (threshold I) and 0.7% (threshold II). For the first time threshold II is not significantly different from zero. At a gain of 0.7% the slope of gains changes from a significant negative level (-0.850) to a level indistinguishable from zero (0.019), perhaps because there is more noise in the region of positive gains. Although the slope is not strictly increasing at the right tail, for consistency reason I still call 0.7% the bottom point of the V shape because the slope of gains experiences a sharp positive change here. The decline of the probability of holding at zero gains is not significant.

Frequent traders holding stocks with good price history (column 3) again have three thresholds estimated at -1.4% (threshold I), 1.8% (threshold II) and 26.4% (threshold III) respectively, all significantly different from zero at 1% significance level. The overall pattern is shown to be V-shaped, very similar to the case of frequent traders as a whole. The bottom point of the V shape (threshold II) is also estimated at 1.8%. Investors in this case are 4% less likely to hold a given stock once zero gains is reached, reflecting the effect of the status quo.

In summary, among frequent traders the bottom point of the V shape is higher in the sample of stocks with good price (1.8%) than with bad history (0.7%), but the distance between them is very small. Frequent traders adjust their reference levels of gains in the same direction of average price realization during the holding period, but not too much. The aggregate pattern of frequent traders closely matches the V-shaped prediction of loss aversion with a fixed reference point.

Table 2.4. A Multi-threshold Model of the Probability of Holding a Stock (Frequent Traders)

	(1) Frequent traders (Overall)	(2) Frequent traders (Bad History)	(3) Frequent traders (Good History)
Estimating Sample			
Threshold I [99% confidence interval]	-0.020*** [-0.045,-0.011]	-0.043*** [-0.059,-0.024]	-0.014*** [-0.037,-0.004]
Threshold II [99% confidence interval]	0.018*** [0.008, 0.041]	0.007 [-0.043,0.038]	0.018*** [0.005,0.047]
Threshold III [99% confidence interval]	0.262*** [0.157, 0.360]	-	0.264*** [0.163,0.370]
Observations	203987	118941	85046
Testing Sample			
β_0 (Slope before threshold I)	0.024* (0.013)	0.003 (0.013)	0.051 (0.035)
$\beta_0+\beta_1$ (Slope between threshold I and II)	-1.954*** (0.341)	-0.850*** (0.184)	-2.582*** (0.620)
$\beta_0+\beta_1+\beta_2$ (Slope between threshold II and III)	0.234*** (0.039)	-0.019 (0.024)	0.271*** (0.039)
$\beta_0+\beta_1+\beta_2+\beta_3$ (Slope after threshold III)	-0.011 (0.014)	-	-0.011 (0.015)
α_1 (Discontinuity at zero)	-0.027*** (0.010)	0.007 (0.010)	-0.040** (0.017)
Adjusted R^2	0.061	0.044	0.079
Observations	202323	117909	84414

Note: This table reports the estimation results of a multi-threshold model that regresses a binary decision to hold a stock or not on gains from that stock and other control variables. The slope of gains is allowed to change at multiple unknown thresholds. I randomly split observations into an estimating sample (to identify the thresholds) and a testing sample (to test the magnitude of the changes). In the estimating sample I use the procedure developed by Hansen (2000) to construct the heteroskedastic-consistent 99% confidence interval for the location of the threshold based on a likelihood ratio test. In the testing sample, because of the sample splitting the test for the slope change is standard. Standard errors clustered by account number are reported in brackets. Control variables include December dummy, holding period, trading frequency, portfolio size, tax rate, income, net wealth, daily trading volume, total shares out, market and the stock's own returns dating back as far as two months. Estimates of control variables are reported in Table 2.4.A of Appendix 2.5.

In Table 2.5 infrequent traders as a whole (column 1) have two thresholds estimated at -4.5% (threshold I) and 10.3% (threshold II), all significantly different from zero at 1% significance level. Before threshold I, one unit increase in gains makes infrequent traders 18.6% more likely to hold a given stock. Between threshold I and threshold II, the slope of gains is significantly negative (-0.586). The slope changes to a level indistinguishable from zero (0.005) after threshold II. Again I call threshold II (10.3%) the bottom point of the V shape even though the slope is not strictly positive at the right tail. Infrequent traders are 7% less likely to hold a stock at zero gains.

Infrequent traders who own stocks with bad price history (column 2) have two thresholds estimated at a gain of -4.3% (threshold I) and 0.9% (threshold II). Similar to the sample of stocks with bad history owned by frequent traders, threshold II is not significantly different from zero. At a gain of 0.9% the slope of gains changes from a significant negative level (-1.324) to a level indistinguishable from zero (-0.038). There is also significant decline in the probability of holding once gains become positive.

Infrequent traders who own stocks with good price history (column 3) have two thresholds estimated at -1.9% (threshold I) and 11.9% (threshold II), both significantly different from zero. The overall pattern is similar to the case of infrequent traders as a whole, with the bottom point of the V shape (threshold II) estimated at a gain of 11.9%. Investors in this case are 8.1% less likely to hold a stock at zero gains.

Table 2.5. A Multi-threshold Model of the Probability of Holding a Stock (Infrequent Traders)

	(1) Infrequent traders (Overall)	(2) Infrequent traders (Bad History)	(3) Infrequent traders (Good History)
Estimating Sample			
Threshold I [99% confidence interval]	-0.045*** [-0.057,-0.034]	-0.043*** [-0.058,-0.025]	-0.019 [-0.061,0.012]
Threshold II [99% confidence interval]	0.103*** [0.073,0.147]	0.009 [-0.027,0.073]	0.119*** [0.078,0.160]
Observations	358094	170384	187710
Testing Sample			
β_0 (Slope before threshold I)	0.186*** (0.011)	0.155*** (0.011)	0.092** (0.039)
$\beta_0+\beta_1$ (Slope between threshold I and II)	-0.586*** (0.043)	-1.324*** (0.171)	-0.487*** (0.044)
$\beta_0+\beta_1+\beta_2$ (Slope after threshold II)	0.005 (0.007)	-0.038 (0.025)	-0.001 (0.008)
α_1 (Discontinuity at zero)	-0.070*** (0.005)	-0.029*** (0.010)	-0.081*** (0.006)
Adjusted R^2	0.156	0.110	0.183
Observations	355281	168379	186902

Note: This table reports the estimation results of a multi-threshold model that regresses a binary decision to hold a stock or not on gains from that stock and other control variables. The slope of gains is allowed to change at multiple unknown thresholds. I randomly split observations into an estimating sample (to identify the thresholds) and a testing sample (to test the magnitude of the changes). In the estimating sample I use the procedure developed by Hansen (2000) to construct the heteroskedastic-consistent 99% confidence interval for the location of the threshold based on a likelihood ratio test. In the testing sample, because of the sample splitting the test for the slope change is standard. Standard errors clustered by account number are reported in brackets. Control variables include December dummy, holding period, trading frequency, portfolio size, tax rate, income, net wealth, daily trading volume, total shares out, market and the stock's own returns dating back as far as two months. Estimates of control variables are reported in Table 2.5.A of Appendix 2.5.

In summary, among infrequent traders the bottom point of the V shape is much higher in the sample of stocks with good price history (11.9%) than stocks with bad history (0.9%). This could result from infrequent traders adjusting their reference points substantially in the same direction of average returns realization in the past. The relationship is not globally V-shaped both because of a positive slope at the left tail and an almost flat relationship at the right tail.

Loss aversion with the reference points defined by expectations alone has a hard time explaining the last observation about patterns at the tails. One possibility comes from the literature on changing risk attitudes in sequential gambles. Thaler and Johnson (1990) find that people generally become more risk averse after prior losses and take more risks after prior gains, mainly because they use a heuristic editing rule to integrate or segregate prior gains and losses. Barberis, Huang, and Santos (2001) incorporate this idea into an asset allocation model with changing risk attitudes. This theory can easily explain the left-tail drop in the probability of holding a stock, but it somehow fails to capture the high level of risk aversion at the right tail. It is also possible that returns in the past two months cannot control for beliefs adequately so there exists momentum beliefs. Such beliefs may explain the observed left-tail pattern, but the implication is again incompatible with the almost flat relationship at the right tail. What's more, learning alone should generate smoother transition rather than the sharp changes observed in the data. Momentum beliefs combined with more extreme movements in the reference points may provide a better account of the patterns at tails. For instance, the adjustment of expectations to both large gains and losses leads to low probability of holding a given stock in these regions. Loss aversion reinforces the effect of momentum beliefs at the left tail but counter-balances it at the right tail. Thus the decreasing in the probability of holding a stock as losses become large is very salient, but at the same time there is almost no change in the probability of holding a stock as gains becomes large.

The left-tail and right-tail changes observed on infrequent traders remain a puzzle. My conjecture is that changes in expectations hence the reference points should play an important role in a satisfactory explanation.

2.5. Discussion and Conclusion

BX (2008, 2009 section III) develop an alternative model of realization utility based on the distinction between paper and realized gains and losses. The optimal solution is characterized by a threshold strategy that makes investors sell the stocks once certain threshold (higher than the purchase price if with transaction cost) is reached. Combined with positive time discounting, realization utility predicts more sales above the purchase price, among a wide range of other predictions. The threshold selling strategy explains better the behavior of infrequent traders than that of frequent traders in this sample. In particular, this strategy has some difficulties in explaining why the probability of holding a stock rises significantly after passing the bottom point of the V shape in the sample of frequent traders. It is also not easy to reconcile the threshold selling strategy with the fact that the aggregate probability of holding a stock begins to decline significantly after gains become larger than -3.9% (estimated threshold I in Table 2.3 column 3) rather than after a slight positive level of gains as a model of realization utility would predict.

I believe that realization utility is an important psychological factor in trading that is complementary to loss aversion in explaining the disposition effect. Incorporating it into any model of reference-dependent preferences should substantially improve the quantitative accuracy of the predictions.

To conclude, this paper relies on aversion to losses relative to a reference point defined by expectations to explain the disposition effect. Loss aversion predicts a V-shaped relationship between the optimal position in a given stock and current gains from that stock, which implies the disposition effect when the reference point is defined expectations but is itself a stronger and novel empirical regularity. Empirical analysis using individual trading records has indeed discovered such V-shaped relationship. The theoretical prediction of loss aversion allows a reasonable econometric identification of the reference point, and the estimates from both the overall sample and heterogeneous subsamples strongly support expectations as the most reasonable candidate. The common assumption in the literature that treats the status quo as the reference point renders prospect theory, the most popular informal explanation, incapable of generating the disposition effect. The theoretical and empirical results from this paper resolve this puzzle by emphasizing the fact that investors' reference points are defined by positive expectations. More careful studies regarding the nature of such reference point and the implications to trading behavior are needed.

Acknowledgements

Chapter 2, in full, has been submitted for publication of the material. Meng, Juanjuan. The dissertation author was the primary author of this material.

Appendix 2.1. Proof of Proposition 2.1

Since the investor assigns any return below the cut-off return $K(x_1)$ a higher weight λ, and any negative return brings negative marginal utility, the expected marginal utility is decreasing in $K(x_1)$ when $K(x_1) < 0$. When current gain is higher than the reference level $g_1 > g_1^{RP}$, only negative future returns can make future wealth equal to the reference point so $K(x_1) < 0$. If $K(x_1)$ is zero, by assumption the expected marginal utility is negative $E(MU(K(x_1) = 0)) < 0$. As $K(x_1)$ approach negative infinity, all the returns are in the gains domain hence weighted equally, so the expected marginal utility is positive. Therefore there exists a cut-off point $K_1 < 0$ that satisfies the first-order condition that the expected marginal utility is zero. The optimal position is then determined by $K(x_1^*) = K_1$.

The case when $g_1 < g_1^{RP}$ or $g_1 = g_1^{RP}$ follows similar analysis.

Appendix 2.2 Dynamic Model (Proof of Proposition 2.2)

The dynamic problem can be written as the following recursive structure:

(2.12) $$V(W_t, W_t^{RP}) = \max_{x_t} E_t(U(W_{t+1} \mid W_t^{RP}) + \beta V(W_{t+1}, W_{t+1}^{RP})).$$

Where

(2.13) $$W_{t+1} = W_t + x_t P_t r_{t+1}.$$

(2.14) $$W_t^{RP} = W_{t-1} + x_{t-1} P_{t-1}((1+r_t^{RP})(1+r_{t+1}^{RP}) - 1).$$

(2.15) $$\overline{P}_{t-1} g_t^{RP} = P_{t-1}((1+r_t^{RP})(1+r_{t+1}^{RP}) - 1).$$

I express the reference point in terms of the reference level of returns in period t (r_t^{RP}) and $t+1$ (r_{t+1}^{RP}) rather than gains from the stock g_t^{RP} for computational convenience. But they are related by equation (2.15).

Let $U_{W_{t+1}} = \dfrac{\partial U(W_{t+1} \mid W_t^{RP})}{\partial W_{t+1}}$, $U_{W_t^{RP}} = \dfrac{\partial U(W_{t+1} \mid W_t^{RP})}{\partial W_t^{RP}}$, $V_{W_{t+1}} = \dfrac{\partial V(W_{t+1}, W_{t+1}^{RP})}{\partial W_{t+1}}$ and $V_{W_{t+1}^{RP}} = \dfrac{\partial V(W_{t+1}, W_{t+1}^{RP})}{\partial W_{t+1}^{RP}}$. The F.O.C is given by equation (2.16).

(2.16) $$E_t(U_{W_{t+1}} r_{t+1} + \beta V_{W_{t+1}} r_{t+1} + \beta V_{W_{t+1}^{RP}}((1+r_{t+1}^{RP})(1+r_{t+2}^{RP}) - 1) = 0$$

Through simple algebra, we know that

(2.17) $$V_{W_t^{RP}} = E_t(U_{W_t^{RP}}),$$

(2.18) $$V_{W_t} = E_t(U_{W_{t+1}} + \beta V_{W_{t+1}} + \beta V_{W_{t+1}^{RP}}),$$

(2.19) $$E_t(U_{W_{t+1}} + U_{W_t^{RP}}) = 0 .$$

Law of iterated expectation equation (2.19) together suggest that

(2.20) $$V_{W_t} = E_t(U_{W_{t+1}} + \beta^{T-t}(V_{W_T} + V_{W_T^{RP}})).$$

In the final decision period T, the indirect value function is $V(W_T, W_T^{RP}) = \max_{x_T} E_T(U(W_{T+1} \mid W_T^{RP})$, which implies both $V_{W_T} = E_T(U_{W_{T+1}})$ and $V_{W_T^{RP}} = E_T(U_{W_T^{RP}})$. According to equation (2.19) the two terms cancel out and we get condition (2.21)

(2.21) $$V_{W_t} = E_t(U_{W_{t+1}})$$

Plug in equation (2.17) and (2.21) into the F.O. C (2.16). After simplifying we get

(2.22) $$E_t(U_{W_{t+1}} r_{t+1} + \beta U_{W_{t+1}^{RP}}((1+r_{t+1}^{RP})(1+r_{t+2}^{RP}) - (1+r_{t+1}))) = 0$$

To have an understanding of the relationship between the optimal position and stock return in period t, I take total derivative with respect to equation (2.22).

Let $K_t = \dfrac{W_t^{RP} - W_t}{x_t P_t} = \dfrac{x_{t-1}}{x_t}\left(\dfrac{(1+r_t^{RP})(1+r_{t+1}^{RP})}{(1+r_t)} - 1\right)$. Instead of looking at $\dfrac{dx_t^*}{dr_t}$, it is computationally more convenient to look at

$$\frac{dr_t}{dx_t^*} = \frac{(1+r_t)}{x_t^*}\left(\frac{(1+r_t)}{(1+r_t^{RP})(1+r_{t+1}^{RP})} - 1\right)$$

(2.23)
$$- \frac{A}{f(K_t^*)\dfrac{x_{t-1}^{*2}}{x_t^{*2}}\dfrac{(1+r_t^{RP})(1+r_{t+1}^{RP})((1+r_t^{RP})(1+r_{t+1}^{RP})-(1+r_t))}{(1+r_t)^3}},$$

where

(2.24)
$$A = \int_{-1}^{+\infty} f(K_{t+1})\frac{((1+r_{t+1}^{RP})(1+r_{t+2}^{RP})-(1+r_{t+1}))^2}{(1+r_{t+1})x_{t+1}^*} f(r_{t+1})dr_{t+1} > 0$$

From equation (2.23) and (2.24) it is easy to see that when $r_t < (1+r_t^{RP})(1+r_{t+1}^{RP}) - 1$, $\frac{dx_t^*}{dr_t} < 0$; when $r_t \geq (1+r_t^{RP})(1+r_{t+1}^{RP}) - 1$, $\frac{dx_t^*}{dr_t} \geq 0$.

Expressing the result in term of gains from the stock g_t, first recall that $P_{t-1}(1+r_t) = \overline{P}_{t-1}(1+g_t)$ and $P_{t-1}(1+r_t^{RP})(1+r_{t+1}^{RP}) = \overline{P}_{t-1}(1+g_t^{RP})$, where $\overline{P}_{t-1}$ is the average purchase price at the end of period $t-1$, g_t is the gains in period t, and g_t^{RP} is the reference level of gains in period t. We also have the same relationship between the optimal position x_t^* and gains from the stock g_t: When $g_t < g_t^{RP}$, the optimal position is decreasing in g_t; when $g_t > g_t^{RP}$, the optimal position is increasing in g_t.

Appendix 2.3. Stochastic Reference Point (Proof of Proposition 2.3)

Assume that g_1^{RP} follows a distribution $h(g_1^{RP})$. The expected marginal utility of adding an extra share is given by equation (2.25).

$$(2.25)\qquad E(MU(K(x_1))) = P_1 \int_{-1}^{+\infty} \Big(\int_{K(x_1)}^{+\infty} r_2 f(r_2) dr_2 + \int_{-1}^{K(x_1)} r_2 f(r_2) dr_2 \Big) h(g_1^{RP}) dg_1^{RP}$$

According to the F.O.C that $E(MU(K(x_1^*))) = 0$, we have

$$(2.26)\qquad \frac{dx_1^*}{dg_1} = -\frac{\int_{-1}^{+\infty} \big(\frac{1+g_1^{RP}}{1+g_1}\big)(g_1^{RP} - g_1) f(K(x_1^*)) h(g_1^{RP}) dg_1^{RP}}{\int_{-1}^{+\infty} \frac{1}{x_1^*}(g_1^{RP} - g_1)^2 f(K(x_1^*)) h(g_1^{RP}) dg_1^{RP}}$$

The denominator is always positive. Thus the relationship between the optimal position and gains depends on the sign of the numerator, which in turn depends on the sign of $(g_1^{RP} - g_1)$ and the properties of the densities $f(.)$ and $h(.)$. When $g_1 = -1$, it follows that $g_1^{RP} - g_1 \geq 0$ for all levels of g_1^{RP}, so $\frac{dx_1^*}{dg_1} \leq 0$. Because $\frac{dx_1^*}{dg_1}$ is continuous in g_1, there exists a lower bound $\underline{g}_1$ such that for $g_1 < \underline{g}_1$ the relationship between optimal position and gains is negative. Similarly, when $g_1 \to +\infty$, it follows that $g_1^{RP} - g_1 < 0$ for all levels of g_1^{RP}, so $\frac{dx_1^*}{dg_1} > 0$. There exists a higher bound $\overline{g}_1$ such that for $g_1 > \overline{g}_1$ the derivative is positive. For $\underline{g}_1 \leq g_1 \leq \overline{g}_1$, the relationship is ambiguous. It depends on the characteristics of $f(.)$ and $h(.)$.

As a relevant example, Let us look at the case of two reference points $(W_1^{RP,L}, p\, ;\, W_1^{RP,H}, 1\text{-}\, p)$ with the application in mind that the low reference point is the status quo and the high one is the mean expectation. Let $g_1^{RP,L}$ and $g_1^{RP,R}$ be the corresponding reference levels of gains for the two reference points, respectively. Through some tedious algebra, we have the following observations:

(i) when $g_1 \leq g_1^{RP,L}$, the optimal position is decreasing in g_1; when $g_1 > g_1^{RP,H}$, the optimal position is increasing in g_1.

(ii) when $g_1^{RP,L} < g_1 < g_1^{RP,H}$ the relationship is ambiguous:

a. Single-trough pattern: if $f(r_2)$ is relatively constant around $r_2 = 0$, there exists a level of gains $\tilde{g}_1$ such that $g_1^{RP,L} < \tilde{g}_1 < g_1^{RP,H}$. The optimal position x_1^* decreases in g_1 in the region $g_1^{RP,L} < g_1 < \tilde{g}_1$, and increases in g_1 in the region $\tilde{g}_1 < g_1 < g_1^{RP,H}$, with the minimum position reached at $g_1 = \tilde{g}_1$.

b. Twin-trough pattern: if $f(r_2)$ is strongly increasing in the region of small negative returns, and decreasing in the region of small positive returns, there exists a level of gains $\hat{g}_1$ such that $g_1^{RP,L} < \hat{g}_1 < g_1^{RP,H}$. The optimal position x_1^* increases in g_1 in the region $g_1^{RP,L} < g_1 < \hat{g}_1$, and decreases in g_1 in the region $\hat{g}_1 < g_1 < g_1^{RP,H}$, with the two local minimum positions reached at $g_1^{RP,L}$ and $g_1^{RP,H}$ respectively.

Appendix 2.4. Data Cleaning Process

1. Construct Trading Records.

In this study I focus on common stocks only. Trading records with short selling are excluded from the sample for simplicity. This is about 2% of the observations. Intraday trades are netted with price aggregated to be the volume weighted average price.

2. Add Position Records.

For each stock held by each account owner, I connect trading data to the end-of-the-month position data in chronological order. I match the first trading record of a stock in the sample with the most recent end-of-the-month position record before the first trading date. Such end-of-the-month position record serves as the initial position to construct the complete trading records for each stock based on the trading data.

Any trading history not starting with an initial position of zero is dropped out of the sample, so I end up with consistent trading histories of stocks whose initial purchase prices are known.

3. Fill in Dates of Hold.

I expand the trading records of each stock by adding dates in which at least one sale is made in the portfolio this stock belongs to. I then obtain daily stock price (either closing price of the date or the average of ask and bid prices if closing price is not available), market returns (S&P), trading volume, number of shares out, and adjustment factor for dividend and split from CRSP. Prices are adjusted for commission, dividend and splits. Commissions for potential sales are assumed to be equal to the average commission incurred when purchasing this stock. To avoid being driven by outliers I winsorize 0.5% of the extreme gains and losses at tails, which restricts gains from each stock to be within (-.788,2.046).

Appendix 2.5. The Estimates of Control Variables in Table 2.3-2.5

Table 2.3.A. A Multi-threshold Model of the Probability of Holding a Stock (Overall Sample), Continued.

	(1)	(2)	(3)
December	-	-0.015*** (0.003)	-0.012*** (0.003)
Holding Days/100	-	0.014*** (0.000)	0.011*** (0.000)
Trading Frequency	-	0.856*** (0.064)	0.850*** (0.066)
Portfolio Size	-	0.002 (0.001)	0.002 (0.001)
Tax	-	0.001*** (0.000)	0.001*** (0.000)
Income/ 10^7	-	0.001 (0.003)	0.002 (0.003)
Net worth/ 10^7	-	0.002* (0.001)	0.001 (0.001)
Total Daily Trading Volume of the Holding Stock/ 10^8	-	-2.323*** (0.110)	-2.401*** (0.124)
Total Outstanding Shares of the Holding Stock/ 10^7	-	0.101*** (0.004)	0.088*** (0.004)
Market Returns 1	-	-	0.749*** (0.112)
Market Returns 2	-	-	0.755*** (0.113)
Market Returns 3	-	-	0.257** (0.109)
Market Returns 4	-	-	0.110 (0.110)
Market Returns 5	-	-	0.110 (0.115)
Market Returns 6~20	-	-	0.170*** (0.043)
Market Returns 21~40	-	-	0.044 (0.035)
Market Returns 41~60	-	-	-0.044 (0.034)
S&P index	-	-	0.000*** (0.000)
Own Returns 1	-	-	-0.286*** (0.023)
Own Returns 2	-	-	-0.155*** (0.032)
Own Returns 3	-	-	-0.118*** (0.036)

Table 2.3.A. A Multi-threshold Model of the Probability of Holding a Stock (Overall Sample), Continued.

Own Returns 4	-	-	-0.172*** (0.022)
Own Returns 5	-	-	-0.178*** (0.019)
Own Returns 6~20	-	-	-0.100*** (0.010)
Own Returns 21~40	-	-	-0.059*** (0.006)
Own Returns 41~60	-	-	-0.047*** (0.005)
Constant	0.788*** (0.001)	0.651*** (0.009)	0.526*** (0.010)
Adjusted R^2	0.018	0.088	0.094
5% and 95% Percentiles of the fitted probability of holding	[0.637,0.783]	[0.545,.932]	[0.530,0.942]
Observations	562,125	561,944	557,604

Table 2.3.A. A Multi-threshold Model of the Probability of Holding a Stock (Frequent Traders), Continued.

	(1) Frequent traders (Overall)	(2) Frequent traders (Bad History)	(3) Frequent traders (Good History)
December	-0.008* (0.004)	-0.028*** (0.005)	0.020*** (0.007)
Holding Days/100	0.013*** (0.001)	0.011*** (0.001)	0.012*** (0.001)
Trading Frequency	0.132** (0.056)	0.188*** (0.059)	0.022 (0.075)
Portfolio Size	0.001** (0.000)	0.001** (0.000)	0.001*** (0.001)
Tax	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Income/ 10^7	-0.063*** (0.019)	-0.060*** (0.020)	-0.071*** (0.024)
Net worth/ 10^7	0.093*** (0.024)	0.064** (0.026)	0.130*** (0.031)
Total Daily Trading Volume of the Holding Stock/ 10^8	-2.791*** (0.157)	-2.320*** (0.174)	-3.031*** (0.181)
Total Outstanding Shares of the Holding Stock/ 10^7	0.081*** (0.006)	0.071*** (0.007)	0.083*** (0.009)
Market Returns 1	0.753*** (0.159)	0.542*** (0.202)	1.099*** (0.243)
Market Returns 2	0.737*** (0.166)	0.860*** (0.195)	0.507** (0.258)
Market Returns 3	0.474*** (0.152)	0.739*** (0.175)	0.074 (0.260)
Market Returns 4	0.093 (0.158)	0.212 (0.190)	0.040 (0.247)
Market Returns 5	0.133 (0.166)	0.289 (0.183)	0.023 (0.279)
Market Returns 6~20	0.137** (0.068)	0.153* (0.078)	0.142 (0.088)
Market Returns 21~40	0.034 (0.056)	0.020 (0.066)	0.055 (0.082)
Market Returns 41~60	-0.043 (0.050)	0.008 (0.061)	-0.106 (0.072)
S&P index	0.000* (0.000)	-0.000 (0.000)	0.000*** (0.000)
Own Returns 1	-0.091** (0.046)	0.208*** (0.048)	-0.416*** (0.069)
Own Returns 2	-0.174*** (0.044)	0.232*** (0.047)	-0.568*** (0.065)
Own Returns 3	-0.121*** (0.038)	0.062* (0.038)	-0.316*** (0.070)

Table 2.3.A. A Multi-threshold Model of the Probability of Holding a Stock (Frequent Traders), Continued.

Own Returns 4	-0.077**	0.009	-0.206***
	(0.037)	(0.037)	(0.043)
Own Returns 5	-0.115***	-0.036	-0.204***
	(0.029)	(0.034)	(0.053)
Own Returns 6~20	-0.045***	-0.050***	-0.021
	(0.012)	(0.013)	(0.016)
Own Returns 21~40	-0.033***	-0.048***	0.006
	(0.009)	(0.010)	(0.015)
Own Returns 41~60	-0.022***	-0.044***	0.028*
	(0.009)	(0.010)	(0.015)
Constant	0.759***	0.797***	0.726***
	(0.021)	(0.021)	(0.027)
Adjusted R^2	0.061	0.044	0.079
5% and 95% Percentiles of the fitted probability of holding	[0.705,0.990]	[0.766, 0.990]	[0.646,0.987]
Observations	202,323	117,909	84,414

Table 2.5.A . A Multi-threshold Model of the Probability of Holding a Stock (Infrequent Traders), Continued.

	(1) Infrequent traders (Overall)	(2) Infrequent traders (Bad History)	(3) Infrequent traders (Good History)
December	-0.011***	-0.078***	0.062***
	(0.003)	(0.004)	(0.004)
Holding Days/100	0.010***	0.006***	0.011***
	(0.001)	(0.001)	(0.001)
Trading Frequency	2.655***	2.621***	2.621***
	(0.271)	(0.216)	(0.343)
Portfolio Size	0.019***	0.016***	0.020***
	(0.003)	(0.003)	(0.004)
Tax	0.000*	0.000	0.001***
	(0.000)	(0.000)	(0.000)
Income/ 10^7	-0.000	0.004	-0.002
	(0.002)	(0.006)	(0.001)
Net worth/ 10^7	0.001	-0.000	0.002**
	(0.001)	(0.002)	(0.001)
Total Daily Trading Volume of the Holding Stock/ 10^8	-1.903***	-1.872***	-1.513***
	(0.124)	(0.099)	(0.209)
Total Outstanding Shares of the Holding Stock/ 10^7	0.094***	0.080***	0.088***
	(0.005)	(0.005)	(0.008)
Market Returns 1	0.820***	0.919***	0.766***
	(0.125)	(0.175)	(0.182)
Market Returns 2	0.853***	1.264***	0.513***
	(0.133)	(0.181)	(0.190)
Market Returns 3	0.298*	0.500**	0.098
	(0.162)	(0.200)	(0.221)
Market Returns 4	0.516***	0.573***	0.518***
	(0.144)	(0.187)	(0.197)
Market Returns 5	0.179	0.339*	0.120
	(0.146)	(0.186)	(0.205)
Market Returns 6~20	0.141**	0.054	0.243***
	(0.064)	(0.070)	(0.080)
Market Returns 21~40	0.063*	0.032	0.126**
	(0.038)	(0.052)	(0.051)
Market Returns 41~60	0.027	-0.004	0.078*
	(0.035)	(0.051)	(0.046)
S&P index	0.000***	0.000	0.000***
	(0.000)	(0.000)	(0.000)
Own Returns 1	-0.310***	-0.001	-0.583***
	(0.026)	(0.026)	(0.052)
Own Returns 2	-0.117***	0.297***	-0.429***
	(0.033)	(0.032)	(0.061)
Own Returns 3	-0.096**	0.096***	-0.238***
	(0.039)	(0.033)	(0.080)

Table 2.5.A . A Multi-threshold Model of the Probability of Holding a Stock (Infrequent Traders), Continued.

Own Returns 4	-0.184*** (0.025)	-0.027 (0.026)	-0.364*** (0.046)
Own Returns 5	-0.167*** (0.021)	-0.056*** (0.021)	-0.280*** (0.048)
Own Returns 6~20	-0.105*** (0.014)	-0.067*** (0.010)	-0.118*** (0.027)
Own Returns 21~40	-0.061*** (0.007)	-0.060*** (0.008)	-0.050*** (0.009)
Own Returns 41~60	-0.049*** (0.006)	-0.068*** (0.008)	-0.019** (0.009)
Constant	0.430*** (0.010)	0.508*** (0.011)	0.363*** (0.012)
Adjusted R^2	0.156	0.110	0.183
5% and 95% Percentiles of the fitted probability of holding	[0.401, 0.959]	[0.518, 0.956]	[0.348,0.956]
Observations	355,281	168,379	186,902

Appendix 2.6. The Residuals of the Probability of Holding a Stock

The residuals of the probability of holding a given stock comes from a linear probability regression where the binary decision to hold a given stock is regressed on the full set of control variables in column 3 of Table 2.3. In general, the residuals look very similar to the raw relationships except that the left-drop in the probability of holding is more severe in the residuals graph.

Figure 2.7. The Residuals of the Probability of Holding a Stock

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. Forthcoming.”Reference Points and Effort Provision." American Economic Review.

Arkes, Hal Richard, David Hirshleifer, Danling Jiang and Sonya Seongyeon LimBai. 2008. “Reference point adaptation: Tests in the domain of security trading” Organizational Behavior and Human Decision Processes, 105 (1): 67-81

Bai, Jushan. 1997. "Estimation of A Change Point in Multiple Regression Models" The Review of Economics and Statistics, 79(4): 551-563.

Bai, Jushan and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes" Econometrica, 66(1): 47-78.

Barber, Brad M. and Terrance Odean. 2000. "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors" The Journal of Finance, LV (2): 773-806.

Barber, Brad M. and Terrance Odean. 2001. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment" The Quarterly Journal of Economics, 116(1): 261-292.

Barberis, Nicholas. 2009. “A Model of Casino Gambling” working paper

Barberis, Nicholas and Ming Huang. 2001. "Mental Accounting, Loss Aversion, and Individual Stock Returns" Journal of Finance, 56(4): 1247-1292.

Barberis, Nicholas, Ming Huang, and Tano Santos. 2001. "Prospect Theory and Asset Prices" The Quarterly Journal of Economics, 116(1): 1-53.

Barberis, Nicholas and Wei Xiong. 2008. "Realization Utility" Working paper.

Barberis, Nicholas and Wei Xiong. 2009. "What Drives the Disposition Effect? An Analysis of A Long-Standing Preference-Based Explanation" Journal of Finance, 64(2): 751-784.

Benartzi, Shlomo and Richard H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle" Quarterly Journal of Economics, 110(1): 73-92.

Camerer, Colin, Linda Babcock, George Loewenstein and Richard Thaler. 1997. “Labor Supply of New York City Cabdrivers: One Day at a Time” Quarterly Journal of Economics, 112(2):407-441.

Card, David and Gordon B. Dahl. 2009. “Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior.” Working paper.

Card, David, Alexandre Mas, and Jesse Rothstein. Forthcoming." Tipping and the Dynamics of Segregation" The Quarterly Journal of Economics.

Coval, Joshua D., and Tyler Shumway. 2005. "Do Behavioral Biases Affect Prices?" Journal of Finance, 60 (1):1-34.

Crawford, Vincent P., and Juanjuan Meng. 2008. “New York City Cabdrivers’ Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income” University of California, San Diego, Department of Economics Discussion Paper 2008-03, http://dss.ucsd.edu/~vcrawfor/cabdriver11.3.pdf.

De Bondt, Werner, and Richard Thaler. 1985. “Does the Stock Market Overrect?” Journal of Finance 40: 793-808.

Ericson, Keith M. Marzilli and Fuster, Andreas. 2009. “Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior” Available at SSRN: http://ssrn.com/abstract=1505121

Fama, Eugene. 1976. “Foundations of Finance” (Basic Books, New York)

Genesove, David and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market" The Quarterly Journal of Economics, 116(4): 1233-1260.

Gill, David and Victoria Prowse. 2009. “A Structural Analysis of Disappointment Aversion in a Real Effort Competition” working paper

Gneezy, Uri. “Updating the Reference Level: Experimental Evidence” in Experimental Business Research, New York: Springer Publisher, 2005, 263-284

Gomes , Francisco J. 2005. "Portfolio Choice and Trading Volume with Loss-Averse Investors" Journal of Business, 78(2): 675-706.

Grinblatt, Mark and Bing Han. 2005. "Prospect Theory, Mental Accounting, and Momentum" Journal of Financial Economics, 78(2): 311-339.

Grinblatt, Mark and Matti Keloharju. 2001. "What make Investors Trade" The Journal of Finance, LVI(2): 586-616.

Hansen, Bruce E. 1996. "Inference when a Nuisance Parameter is Not Identified Under the Null Hypothesis" Econometrica, 64(2): 413-430.

Hansen, Bruce E. 2000a. "Sample Splitting and Threshold Estimation" Econometrica, 68(May): 575-603.

Hansen, Bruce E. 2000b. "Testing for Structural Change in Conditional Models" Journal of Econometrics, 97: 93-115.

Hens, Thorsten and Martin Vlcek. 2005. "Does Prospect Theory Explain the Disposition Effect?" Working paper.

Health, Chip, Steven Huddart and Mark Lang. 1999. “Psychological Factors and Stock Option Exercise” The Quarterly Journal of Economics, 114 (2): 601-627

Ivković, Zoran; James Poterba and Scott Weisbenner. 2005. “Tax-Motivated Trading by Individual Investors” The American Economic Review, 95(5):1605-1630

Kőszegi, Botond and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences" The Quarterly Journal of Economics, CXXI (4).

Kőszegi, Botond and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes" American Economic Review, 97(4): 1047-1073.

Kőszegi, Botond and Matthew Rabin. 2009. "Reference-Dependent Consumption Plans" American Economic Review, 3(909): 936.

Kahneman, Daniel, J. L. Knetsch, and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem" Journal of Political Economy, 98(6): 1325-1348.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk" Econometrica, 47(2): 263-292.

Odean, Terrance. 1998. "Are Investors Reluctant to Realize their Losses?" The Journal of Finance, 53(5): 1775-1798.

Odean, Terrance. 1999. "Do Investors Trade Too Much?" The American Economic Review, 89(5): 1279-1298.

Post, Thierry, Martijn J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal Or no Deal? Decision Making Under Risk in a Large-Payoff Game show" American Economic Review, 98(1): 38-71.

Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem” Econometrica, 68(5): 1281-1292.

Rabin, Matthew. 2002. "Inference by Believers in the Law of Small Numbers" The Quarterly Journal of Economics, 117(3): 775-816.

Shefrin, Hersh and Meir Statman. 1985. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence" The Journal of Finance, 40(3): 777-790.

Sydnor, Justin. 2009. “(Over)insuring Modest Risks” working paper.

Thaler, Richard H. and Eric J. Johnson. 1990. "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice" Management Science, 36(6): 643-660.

Tversky, Amos and Daniel Kahneman. 1971. "Belief in the Law of Small Numbers" Psychological Bulletin, 76(2): 105-110.

Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty" Journal of Risk and Uncertainty, 5: 297-323.

Weber, Martin and Colin F. Camerer. 1998. "The Disposition Effect in Securities Trading: An Experimental Analysis" Journal of Economic Behavior & Organization, 33: 167-184.

Chapter 3

Social Distance, Interpersonal Interaction and Impersonal Exchange

3.1. Introduction

Whether buyers and sellers rely on personal relationships or an impersonal market to exchange goods and services is a central question underlying economic development. Substantial theoretical and empirical studies have been devoted to discuss the differences between these two forms of exchange from the perspective of social dilemma and contract enforcement (e.g. Greif, 1994; Kranton, 1996a; Johnson, McMillan, and Woodruff, 2002; Dixit, 2003; Kimbrough, Smith, and Wilson, 2008). In the theoretical literature, exchange between buyers and sellers is normally modeled as a prisoner's dilemma game, in which mutual defection is Nash equilibrium and cooperation can be maintained only through long-run repetition or external contract enforcement. In the absence of a formal legal system to enforce exchange contracts, it has been shown that the threat to terminate an ongoing personal relationship can be an effective enforcement mechanism under one of the following circumstances: information transmission and reputation building within socially close group (Greif, 1994; Dixit, 2003; Bowles and Gintis, 2004), frictions in searching for new match (Kranton, 1996a; Sobel, 2006), higher rewards for cooperation that increases the cost of defection (Greif, 1994), or costly signaling in the beginning of the relationship that increases the cost of establishing new relationship (Kranton, 1996b; Carmichaeland MacLeod, 1997; Leeson, 2008). While earlier studies exogenously assume one of the two exchanges in a given society, Kranton (1996a) and Dixit (2003) provide insights about the endogenous emergence of different modes of exchange. Based on different channels, both studies show that either personal relationships or an impersonal market can be effective in enforcing exchange contracts. However, depending on the fundamentals, only one of them will emerge in equilibrium.[59]

Empirical studies nonetheless suggest cases that the two forms of exchange often coexist, even in countries with well-established legal system (Macaulay, 1963). Further, the two forms of exchange not only differ in terms of enforcement mechanism but also exchange surplus. For instance, in a survey study, Johnson, McMillan, and Woodruff (2002) find that firms in post-communist countries are often engaged in exchanges enforced both by courts and personal relationships, but that different kinds of exchange is enforced in different ways: Courts are used to enforce simple exchange between unfamiliar parties, but relationships support more complex exchange. They also find that the two forms of exchange foster different values: Firms offer higher level of trade credit to those who they have relationships with, and the specific amount varies with the degree of familiarity and whether the customer is a family member or friend. To reciprocate, customers who are in relational contracts pay more percentage

[59] In Kranton (1996a), a larger formal market generates higher outside options for those in the personal relationships, attracting more people to participate the formal market. Therefore, depending on the initial proportion of people engaging in different forms of exchange, in equilibrium traders will converge to one form of the exchange or the other all together. In Dixit (2003), local information transmission about traders' past behavior can enforcement contracts only up to a certain size of the economy, beyond which a formal legal system will emerge to enforce contracts.

points of the bill on credit. On the contrary, exchange enforced by courts that is impersonal in nature is often associated with lower level of credit extended. There is little internal value to such relationship because firms easily switch to suppliers with lower price.[60]

Although existing models have generated many important insights regarding the two forms of exchange and their implications to economic development, most of them cannot account for these empirical regularities simultaneously. This paper adds to Kranton's (1996b) and Dixit's (2003) analyses by allowing a social component in surplus that depends on the mode of exchange, with interpersonal interactions yielding benefits that decline with social distance, but impersonal exchange yielding benefits independent of social distance. This social component constitutes one major difference between personal relationships and an impersonal market that is not thoroughly investigated before.[61] It can be monetary in nature, as different social distances facilitate different levels of information transmission and foster different degrees of reciprocity, or trust and trustworthiness that generally lead to different payoffs in a social dilemma situation.[62] [63] It can also be purely preferences that are independent of

---

[60] Also see Macaulay (1963), Akerlof (1982), and Hart and Moore (2008) for similar argument.

[61] Dixit (2003) and Bowles and Gintis (2004) also imbed the idea of social distance in their models, but mainly as a factor that affects information transmission but not that generates differential payoffs.

[62] Kranton (1996a) and Dixit (2003) assume that there are additional benefits from trading with (socially or geometrically) distant party, a classical assumption in trade theory. The discussion in this paper holds this aspect of the matching quality constant by assuming homogenous goods. In many potential applications of my model, such as doctor-patient or employer-worker relationship, this aspect of the trading benefits may not play a crucial role. Adding these benefits to the model, impersonal market exchange generally gains advantage compared to personal relationships because the former, by being neutral to social distance, facilitates trades between distant parties.

[63] In this approach I follow Akerlof (1982), who characterizes labor contracts as partial gift exchange where workers have reciprocal preferences to firms who pay above-average wage; and Benjamin (2008), who investigates the conditions under which the most general form of social preferences can lead to efficiency in a one-sided bilateral exchange relationship; but with more emphasis on the effects of social distance, which have crucial effects in developing economics.

enforcement problem, such as friendship and degree of comfort in interaction. On the contrary, an impersonal exchange ensures equality across social distance. For instance, the development of a formal market is normally accompanied by the establishment of institutions governed by impersonal rules and regulations that effectively prevent differential treatments on customers. Anonymous interactions, such as online trading and ATM machine, also facilitate such impersonal exchange. Depending on social distance, interpersonal interaction is assumed to generate surplus higher or lower than impersonal exchange, reflecting the idea that interaction between socially too distant parties can be detrimental.

Given the reduced-form assumption of social surplus, my paper investigates how sellers sort into personal or impersonal sector in equilibrium, adapting Salop's (1979) classical model of monopolistic competition to reflect social distance. My model generates cases in which personal and impersonal sectors coexist in equilibrium. This coexistence result follows whenever a seller and the farthest buyer she sells to are socially neither too close nor too far, because the competitive advantage of personal exchange is inversely related to the degree of social heterogeneity. If the average social distance in a given society is proportional to the size of the economy, then my model can replicate the classical result that as the economy grows, anonymous market exchange gains increasing advantage (e.g. Greif, 1994; Kranton, 1996a; Dixit, 2003). However, with non-trivial social structure my model leads to richer predictions. For instance, even when the size of the economy is large, as long as the society is relatively homogenous, an impersonal sector may still not emerge out of personal relationships.

My model also yields interesting welfare and inequality results. While it is generally observed that relying on personal relationships creates barrier of entry (Akerlof, 1982; Akerlof and Yellen, 1990; Johnson, McMillan, and Woodruff, 2002), the resulting inequality, including a comparison of the degrees of inequality under the two forms of exchange, has been under-emphasized in the literature. My model suggests that compared to going to the personal sector, a seller's decision to go to the impersonal sector improves aggregate welfare and degree of equality among buyers in most cases. However, the opposite is true when a given personal seller is socially relatively close to the farthest buyer she sells to. In this case, personal sellers have competitive advantages over impersonal sellers; therefore they can afford to charge a higher price when competing with an impersonal seller relative to the case when they compete with another personal seller. Since price constitutes the component of payoffs that is equal for every buyer, a higher price leads to not only lower welfare but also high inequality measure among buyers. This surprising result suggests that in transitional economy where impersonal market exchange barely starts but personal relationships are still deeply rooted in the society, the establishment of an impersonal sector may lead to unintended consequences.

The rest of the paper is organized as follows. Section 3.2 presents a model of buyers and sellers with exchange surplus defined over social distance. The cases of a single seller and multiple sellers are analyzed. Section 3.3 focuses on comparing the equilibrium outcomes with no impersonal seller and with only one impersonal seller. Welfare and inequality implications among buyers in the two cases are also discussed. Section 3.4 comments on interesting extensions of the model and concludes the paper.

3.2. The Model

3.2.1. Set-up

My model adapts Salop's (1979) classical model of monopolistic competition with outside goods to reflect social distance. While in Salop's (1979) model the distance between buyers and sellers is interpreted as tastes for different brands, in my model it represents social distance that entails differential social surplus. Further, although Salop takes the outside market as given exogenously, sellers' decision on whether to engage in impersonal exchange is my central focus.

• Social surplus

A large amount of laboratory evidence demonstrates that, in the absence of long-run interactions that might foster reputational effects, altruism, reciprocity, and levels of trust and trustworthiness may lead to efficiency in a one-shot social dilemma situation. These factors are shown to be inversely related to social distance. (e.g. Hoffman, McCabe, and Smith, 1996; Bohnet, and Frey,1999; Bernhard, Fehr, and Fischbacher, 2006; Buchan, Johnson, and Croson, 2006; Charness, Haruvy, and Sonsino, 2007; Charness, and Gneezy, 2008).[64] Although existing studies model prisoner's dilemma game explicitly, to focus on the comparison between different forms of exchange, I directly introduce social surplus, modeling the effects of social distance via a reduced form. Social surplus can also be extended broadly to include intrinsic matching qualities in the forms of communication costs, degree of comfort, and friendship, even in the absence of enforcement problem. Akerlof (1997) and

---

[64] Experimenters vary the degree of social distance by running treatments where they reveal information about the opponents' last name, photos, or even allow communication (face-to-face or online) between players before playing the game.

Akerlof and Kranton (2000) present extensive evidence on how this aspect of preferences affects social choices such as education and occupation.

Social structure is represented by a circle with circumference $2h$. A point on the circle might represent a variety of social characteristics such as gender, race, religion, cultural background, education, or occupation, but simplified into a one-dimensional measure for tractability. The length of the arc between any two points s is the social distance between the two social positions. On the circle, the maximum possible social distance is h.

Denote the social component of the exchange surplus by $-k(s-\bar{s})$. k governs the degree of interpersonal interaction. For simplicity, assume that k can only take two possible values.[65] When $k=0$, the exchange is impersonal, and the seller is located outside of the circle in an impersonal sector. When $k=\bar{k}$, the exchange is carried out with social surplus and the seller has a social position on the circle. $\bar{k}>0$ is the degree of interpersonal interaction inherent in the exchange process of a given goods or service. For instance, a doctor-patient relationship naturally involves higher degree of interpersonal interaction than a customer-cashier interaction. $\bar{s}$ is a neutral level of social distance, the existence of which allows for both positive and negative social surplus. The positive-negative distinction is relative to the zero social surplus in an neutral and impersonal exchange. In the presence of negative social surplus, the neutrality of an impersonal exchange is advantageous. In real-life scenarios, an

[65] If we allow k to vary continuously between $[0,\bar{k}]$, it turns out that in most cases, a seller's optimal choice of k still lies on one of the two end points. Exception occurs when there are multiple sellers and the number of sellers N is smaller than $\frac{3h}{2\bar{s}}-8$.

interpersonal interaction between socially very distant parties is more likely to involve distrust, negative reciprocity, hostage, and miscommunication that yield benefits lower than an impersonal transaction. For instance, empirical studies provide evidence that social stigma affects the take-up rate of welfare programs among low-income group in the United States (Moffitt, 1983; Currier and Grogger, 2001; Stuber and Kronebusch, 2004). Akerlof (1997) also comments on the possibility of a social exchange generating negative benefits (1997, p.1011):

> Such negative benefits from social interaction may also reflect reality since not all social exchange contributes positively to utility and fear of negative sanctions, due, for example, to jealousy and envy, are potentially as important a motive for conformity as the desire for the positive benefits of love and friendship.

• Buyers

Assume that l buyers are uniformly distributed around the circle. Each of them demands either one unit of the goods, which generates consumption utility u, or no goods at all and gets zero consumption utility. When the buyer pays price P to purchase one unit of goods from an impersonal seller, her utility is simply $u-P$. When she purchases from a personal seller that is s away from her, however, her utility consists of the standard economic payoff as well as social surplus, as shown in equation (3.1).

$$U^b = (u-P) - \overline{k}(s-\overline{s}) \tag{3.1}$$

• Sellers

Assume that the production technology demonstrates constant marginal cost c, and there is no fixed cost. Sellers decide on whether to go to the impersonal sector ($k = 0$) or personal sector ($k = \bar{k}$). A seller in the personal sector is located on the circle and can sell to buyers in either direction. Since only the distance matters, I ignore direction and focus on distance. If the personal seller sells to buyers within a distance of s on one side, the total units sold are $\frac{l}{2h}s$. For simplicity, I focus on s as representing the quantity sold in the following discussion. If the seller maintains personal relationships with buyers, she also receives the summation of the social surplus from interacting with every buyer within a distance of s:

(3.2) $$-\bar{k}\frac{l}{2h}\int_0^s (a-\bar{s})\,da = -\bar{k}\frac{l}{2h}\left(\frac{1}{2}s^2 - s\bar{s}\right).$$

3.2.2. One Seller

This section analyzes the case of a single seller to demonstrate the implications of adding social surplus, in preparation for the case of multiple sellers.

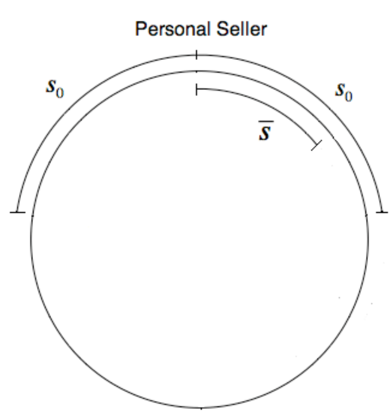


Figure 3.1. The Economy with a Single Seller

Let subscript 0 denote the strategy of a monopolistic seller. s_0 is the social distance of the farthest buyer supplied by the single seller, also the indirect measure of quantity sold, as illustrated by Figure 3.1. The neutral social distance $\bar{s}$ may or may not be smaller than s_0 (Figure 3.1 illustrates the case $\bar{s} < s_0$). It is easy to see that the demand curve of a personal seller is determined by this marginal buyer who is indifferent from purchasing one unit of the goods or not.

$$P_0 = u - \bar{k}(s_0 - \bar{s}) \tag{3.3}$$

The single seller makes two decisions: She first chooses whether to become impersonal or not, i.e. $k_0 = 0$ or $k_0 = \bar{k}$, and then maximize utility given the chosen form of exchange. When $k_0 = 0$, the problem is reduced to a standard monopoly problem. The monopolistic power enables the seller to set $P_0 = u$ and $s_0 = h$.

When $k_0 = \bar{k}$, the single personal seller's decision problem is given by

$$\max_{s_0} U_0^s = \frac{l}{h}((P_0 - c)s_0 - \bar{k}(\frac{1}{2}s_0^2 - s_0\bar{s})), \tag{3.4}$$

subject to demand given in equation (3.3), capacity constraint $0 \le s_0 \le h$ and participation constraint $U_0^s > 0$. Equation (3.5) gives the optimal interior choice.

$$\hat{s}_0 = \frac{u - c}{3\bar{k}} + \frac{2}{3}\bar{s} \tag{3.5}$$

It is clear that the existence of social surplus restricts expansion of sale in many cases because high $\bar{k}$ and low $\bar{s}$ make the personal seller only sell to socially close buyers. The intrinsic economic surplus $u - c$, on the contrary, facilitates expanding of sale to buyers who are socially distant from the personal seller.

Proposition 3.1. It is weakly optimal to choose $k_0 = 0$ when $\bar{k} \le k_s = \dfrac{3(u-c)}{4\bar{s}^2}(h - \dfrac{2}{3}\bar{s} + \sqrt{h^2 - \dfrac{4}{3}\bar{s}h})$.

Proof. See Appendix 3.1.

Proposition 3.1 can be understood through the effects of k_0 on economic profit and social surplus. The personal seller's economic profit depends on social surplus indirectly through the marginal buyer's willingness to pay. When $s_0 \le \bar{s}$ the marginal buyer enjoys positive social surplus, and so $k_0 = 0$ lowers her willingness to pay. The case is the opposite when $s_0 > \bar{s}$. The personal seller's total social surplus depends on s_0 in a non-monotonic way. It is increasing in s_0 when $0 \le s_0 \le \bar{s}$ because of the accumulation of positive social surplus from interacting with buyers in this range, but it is decreasing when s_0 goes larger. When $s_0 > 2\bar{s}$ the overall social surplus becomes negative. Thus $k_0 = 0$ increases the seller's social surplus as long as $s_0 > 2\bar{s}$, and vise visa.

When $u - c \ge 0$, there is positive economic surplus from the exchange. $k_0 = 0$ facilitates expansion of sale, but at the same time the personal seller may suffer in social surplus from interacting with socially distant buyers. As a result, the optimal choice depends on the comparison between the two end points 0 and $\bar{k}$. A low neutral social distance $\bar{s}$ relative to the maximum distance h implies that expanding sale too far is likely to be very costly unless the seller is neutral to social distance. If the degree of interpersonal interaction $\overline{k}$ is lower than the threshold k_s, then positive social surplus generated by personal relationships is low, so the economic profit from sale expansion motives the seller to engage in impersonal exchange. With some calculations it is also clear that the cutoff point k_s is increasing in $u-c$, decreasing in $\overline{s}$ and increasing in h, which are also consistent with the above intuition.

When $u-c<0$, there is no economic surplus inherent in the exchange, so no impersonal exchange is expected. There may be additional benefits from personal relationships as long as the personal seller only interacts with socially close buyers. Therefore $k_s<0$, and it is always optimal to have $k_0=\overline{k}$.

3.2.3. Multiple Sellers

Competition between personal and impersonal sellers poses additional complication. Depending the social distance of the marginal buyer, becoming socially neutral may or may not be an advantage. Figure 3.2 illustrates a possible case of the economy with multiple sellers, in which different personal sellers have different market shares.

Assume that N sellers play a two-stage game. In the first stage, they simultaneously choose the modes of exchange $k_i \in \{0,\overline{k}\}$, $i=1,2,...,N$. Suppose that there are N_p personal sellers with $k_n=\overline{k}$, $n=1,2,...,N_p$ and $N-N_p$ impersonal sellers with $k_m=0$, $m=N_p+1,...,N$. In the second stage they choose quantity of goods s_i to maximize utility.

This paper focuses on symmetric pure strategy subgame perfect Nash equilibrium (SSPNE) where each active seller has positive utility.[66] Symmetry means that all personal sellers choose the same s_n , $n = 1,2,...,N_p$, and all impersonal sellers choose the same $s_m, m = N_p + 1,...,N$. It turns out that symmetry on personal sellers is the feature of the unique pure strategy SPNE, while symmetry on impersonal sellers is exogenously imposed as a restrictive equilibrium selection criterion. It also turns out that only the aggregate behavior of the impersonal sector matters, not the number of impersonal sellers. So one can think of the impersonal sector as a big formal institution, where if a seller chooses $k_m = 0$, she is employed by and shares the profit of the institution with other impersonal sellers.

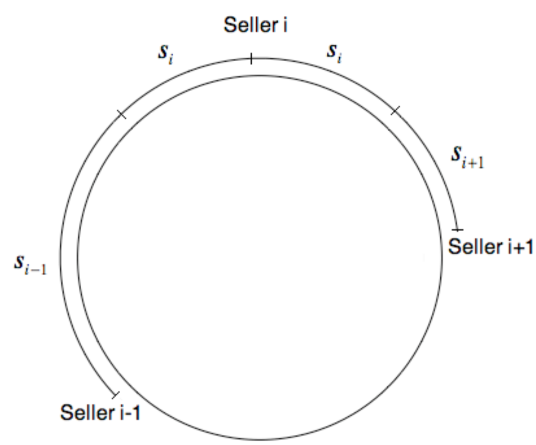


Figure 3.2. The Economy with Multiple Sellers

[66] This positive utility participation constraint suggests that if a seller is indifferent between actively participating in the market and not participating at all, I focus on equilibrium in which she does not participate.

• The second-stage game.

The second-stage maximization problems for personal seller n and impersonal seller m are given by equation (3.6) and (3.7) respectively,

(3.6) $$\underset{s_n}{Max}\ U_n^s = \frac{l}{h}((P_n - c)s_n - \bar{k}(\frac{1}{2}s_n^2 - \bar{s}s_n))$$

(3.7) $$\underset{s_m}{Max}\ U_m^s = \frac{l}{h}((P_m - c)s_m$$

subjects to demand for each type of sellers, participation constraints $U_n^s > 0$, $U_m^s > 0$ and capacity constraint in equation (3.8)

(3.8) $$\sum_{n=1}^{N_p} s_n + \sum_{m=N_p+1}^{N-N_p} s_m = h.$$

Lemma 3.1. It is optimal for personal seller n, $n = 1,2,...,N_p$ to exchange with buyers within equal distance in both directions.

Proof. See Appendix 3.2.

Lemma 3.2. In any pure strategy SPNE, if a personal seller n makes positive utility $U_n^s > 0$, then it is impossible that she competes with personal sellers other than her immediate neighbors.

Proof. See Appendix 3.2.

Lemma 3.3. In any symmetric pure strategy SPNE, if an impersonal seller m makes positive profit $U_m^s > 0$, then all personal sellers directly compete with impersonal sellers.

Proof. See Appendix 3.2.

Lemma 3.2 and 3.3 together imply that in any symmetric pure strategy SPNE where sellers' participation constraints are satisfied, if a given personal seller faces competition from other sellers, then she either directly competes with her neighboring personal sellers, or with impersonal sellers.

In the first case, assume that personal seller n competes with her immediate neighboring personal seller $n+1$ on the right and personal seller $n-1$ on the left. There exist two marginal buyers on each side of personal seller n that lead to the following indifferent condition, $n = 1,2,...,N$

(3.9) $$u - P_{n-1} - \bar{k}(s_{n-1} - \bar{s}) = u - P_n - \bar{k}(s_n - \bar{s}) = u - P_{n+1} - \bar{k}(s_{n+1} - \bar{s}).$$

Define $B_n = P_n + \bar{k}(s_n - \bar{s})$. Equation (3.9) suggests that in equilibrium all the indifferent marginal buyers obtain the same utility hence $B_1 = B_2 = ... = B_N = B$. Price is then given by

(3.10) $$P_n = B - \bar{k}(s_n - \bar{s}).$$

In the second case, assume that personal seller n competes with impersonal seller m, Then the marginal different sellers satisfy equation (3.11) for $n = 1,2,...,N_p$, and $m = N_p + 1,...,N$

(3.11) $$u - P_n - \bar{k}(s_n - \bar{s}) = u - P_m$$

For impersonal sellers, $B_m = P_m$. Imposing symmetry on impersonal sellers' strategy, we also have $B_1 = B_2 = ... = B_N = B$.

From seller n's perspective, B measures the indifferent marginal buyer's total (economic and social) cost of purchasing from the neighboring personal sellers or an impersonal seller, which is also the maximum willingness to pay for seller n's goods before considering social surplus. B directly reflects the degree of competition among sellers. It replaces u, the inherent consumption utility of one unit of the goods, as the major determinant of buyers' willingness to pay. However, for personal sellers to compete with each other rather than maintain a complete monopolistic power over a small territory, we need to impose the condition that $B \le u$. When $B > u$ the total cost of purchasing from any seller outweighs the inherent consumption utility, making indifferent marginal buyers not willing to purchase from any seller. The problem is then reduced to the case of a single seller. To make the case interesting I assume $B \le u$ throughout the following analysis.

Proposition 3.2. In the second-stage of the game,

- When $N_p = N$, it is optimal to set $\hat{s}_{n,N} = \frac{h}{N}$, $n = 1,2,...,N$. The corresponding price is $\hat{P}_{n,N} = \bar{k}(\frac{2h}{N} - \bar{s}) + c$, and each personal seller's utility is $U^s_{n,N} = \bar{k}\frac{3lh}{2N^2}$.

- When $N_p = 0$, in a symmetric pure strategy equilibrium, $\hat{s}_{m,0} = \frac{h}{N}$, $m = 1,2,...,N$. In the absence of competition among impersonal sellers, $\hat{P}_{m,0} = u$, and each impersonal seller's utility is $U^s_{m,0} = \frac{l}{N}(u - c)$.

Proof. See Appendix 3.3.

Proposition 3.2 gives the optimal solution when the market is totally dominated by personal sellers or by impersonal sellers. In the former case, the symmetry in personal sellers' strategies is endogenously determined. It can be shown that the cost of purchasing of marginal buyers is $B = \bar{k}(\frac{3h}{N} - 2\bar{s}) + c$. B is naturally decreasing in the number of sellers N, because more sellers bring more competition. It is also decreasing in $\bar{s}$ and increasing in h because low neutral level of social difference relative to the maximum social distance brings more negative social surplus, leading to higher total cost of being involved in any exchange. In the latter case, the symmetry in quantity is exogenously imposed. Without competition from impersonal seller, the simplifying assumption that each buyer either demands one unit or none of the goods leads to indeterminacy in price. The model is therefore not totally informative when $N_p = 0$. If we treat the impersonal sector as a big formal institution and each impersonal seller as its employee, which is normally the case in reality, then the monopolistic power leads to a price equivalent to the consumption utility u.[67]

Proposition 3.3. For each $0 < N_p < N$, a unique symmetric pure strategy Nash equilibrium exists in the second-stage game, in which

(3.12) $$\widehat{s}_{n,N_p} = \frac{h}{2N_p} + \frac{1}{3}\bar{s}$$

(3.13) $$\widehat{s}_{m,N_p} = \begin{cases} \dfrac{3h - 2\bar{s}N_p}{6(N - N_p)}, \ if \ N_p < \dfrac{3h}{2\bar{s}} \\ 0, \ if \ N_p \geq \dfrac{3h}{2\bar{s}} \end{cases}.$$

[67] If instead we want to have competition among impersonal sellers, then price equal the marginal cost c, and none of impersonal sellers makes any profit.

The corresponding price and utility for each seller type are given by

$$\widehat{P}_{n,N_p} = \overline{k}\left(\frac{h}{N_p} - \frac{1}{3}\overline{s}\right) + c \tag{3.14}$$

$$U^s_{n,N_p} = \overline{k}\,\frac{3l}{2h}\left(\frac{h}{2N_p} + \frac{1}{3}\overline{s}\right)^2 \tag{3.15}$$

$$\widehat{P}_{m,N_p} = B = \overline{k}\left(\frac{3h}{2N_p} - \overline{s}\right) + c \;\; if \;\; N_p < \frac{3h}{2\overline{s}} \tag{3.16}$$

$$U^s_{m,N_p} = \begin{cases} \overline{k}\,\dfrac{l}{12h(N-N_p)N_p}(3h - 2N_p\overline{s})^2, \; if \;\; N_p < \dfrac{3h}{2\overline{s}} \\ 0, \; if \;\; N_p \geq \dfrac{3h}{2\overline{s}} \end{cases} \tag{3.17}$$

Proof. See Appendix 3.3.

It is interesting to see that the quantity and price of personal sellers do not depend on the number of impersonal sellers. Neither does the price of impersonal sellers. This is the case because the homogeneity of impersonal sellers makes only the existence of impersonal sector itself matter, but not the number of impersonal sellers. Thus it is without loss of generality to treat the impersonal sector as a whole and focus on the effect of its presence or absence, as I shall do later.

As the number of personal sellers N_p decreases, the quantities produced by each personal seller and the impersonal sector expands, together with price charged by each sector. This is because N_p determines the number of effective competitors (the impersonal sector as a whole is counted as only one effective competitor), and a reduction in N_p reduces the degree of competition. The neutral level of social distance

$\bar{s}$ is positively related to the quantity produced by a given personal seller, but negative related to the quantity produced by the impersonal sector, because higher $\bar{s}$ leads to positive social surplus for more interactions, making the personal sector more advantageous. Its increase also leads to reduction in price charged by both sectors, because the general cost of purchasing of marginal buyers is lower in the system. Finally, the increase in h implies more demand, so price and quantity of each sector naturally increase. Notice also that $\widehat{s}_{m,N_p} > 0$ and $\widehat{P}_{m,N_p} > c$ only when $N_p < \dfrac{3h}{2\bar{s}}$. For $N_p \geq \dfrac{3h}{2\bar{s}}$ competition is so intensive that no quantity of goods produced is profitable for impersonal sellers.

• The first-stage game.

In equilibrium, no personal seller wants go to the impersonal sector, and vise versa. Therefore, the necessary and sufficient conditions for N_p^* to be in the equilibrium strategy are

(3.18) $$U^s_{n,N_p^*} - U^s_{m,N_p^*-1} \geq 0$$

(3.19) $$U^s_{m,N_p^*} - U^s_{n,N_p^*+1} \geq 0$$

Because of the homogeneity of impersonal sellers, I first characterize the conditions for the existence of the impersonal sector ($N_p < N$), and then analyze the equilibrium number of personal sellers N_p.

Proposition 3.4. Let N_1 be the solution to

$$\bar{k}\frac{3lh}{2N_1^2} - \bar{k}\frac{l}{12h(N_1-1)}(3h-2(N_1-1)\bar{s})^2 = 0 . \quad (3.20)$$

When $N < \min(\frac{3h}{2\bar{s}}+1, N_1)$ and $\bar{s} < \frac{3}{2}(1-\frac{1}{\sqrt{2}})h$, any symmetric pure strategy SPNE involves $N_p < N$.

Proof. See Appendix 3.4.

Proposition 3.4 characterizes the conditions that the impersonal sector will emerge in equilibrium. $N < \frac{3h}{2\bar{s}}+1$ guarantees that the profit of impersonal sellers, if any, is positive. $N < N_1$ implies that it is beneficial for a personal seller to become the only impersonal seller. Lower $\bar{s}$ relative to h and lower N each contributes to the presence of the impersonal sector, because lower tolerance for social distance makes interpersonal interaction between socially distant buyers and sellers less pleasant, and low N increases the incidences of such interaction. Surprisingly $\bar{k}$ does not affect the equilibrium outcome as in the case of a single seller, which confirms the insight that competition is driving the major results in the case of multiple sellers. This analysis is confirmed by proposition 3.5, in which the equilibrium number of personal sellers N_p^* also responses to the changes in parameters in the same direction.

Proposition 3.5. Given the existence of the impersonal sector, let $0 < N_2 < N$ be the solution to $U^s_{n,N_2} - U^s_{m,N_2-1} = 0$. The equilibrium number of personal sellers N_p^*

is given by the largest integer that satisfies $N_p \le N_2$. If N_2 is an integer itself, then either N_2 or $N_2 = 1$ can be in an equilibrium strategy. Further, N_p^* is increasing in $\overline{s}$ and N.

Proof. See Appendix 3.5.

3.3. A Comparison of the Cases with a Single Impersonal Seller and with No Impersonal Seller.

This section focuses on comparing two particular equilibrium outcomes, the one with no active impersonal sector $(N_p^* = N)$ and the one with a single impersonal seller $(N_p^* = N - 1)$. This comparison is important because it reflects the most crucial step of the transition experienced in many developing economies. Gaining more insights from the emergence of an impersonal sector out of personal relationships facilitates further understanding of the quantitative change in the relative market shares of two sectors.

3.3.1. Sellers' Utility

Let N_3 be the solution to $U^s_{n,N_3-1} - U^s_{m,N_3-2} = 0$. It can be easily shown that any integer $N \ge N_3$ will satisfy condition (3.18) and any $N \le N_1$ will satisfy condition (3.19) for $N_p = N - 1$. Proposition 3.4 also suggests that $N_p^* = N$ as long as $N \ge N_1$. It is also interesting to compare personal sellers' utility when there is one or no impersonal seller. Similarly, let N_4 be the solution to $U^s_{n,N_4} - U^s_{n,N_4-1} = 0$. Some calculations suggest that when $N < N_4$, personal sellers' utility is lower under competition with the impersonal sector.

Figure 3.3 graphs the simulated values of N_1, N_3 and N_4 against different values of $\overline{s}$. The parameters are chosen to be $\{h = 0.5, k = 1, u = 2, c = 0, l = 10\}$. Such parameter values guarantee that $B < u$. The chosen range $0 < \overline{s} < 0.2$ can demonstrate all relevant cases. For any pair $\{N, \overline{s}\}$ on the right of the curve representing $N = \frac{3h}{2\overline{s}} + 1$ the utility of any impersonal seller is zero. Any $\{N, \overline{s}\}$ on the right of N_1 entails the equilibrium in which there is no impersonal seller. Similarly, any $\{N, \overline{s}\}$ located between the curves representing N_1 and N_3 leads to the equilibrium with one impersonal seller. The simulation confirms the analytical result that small N and $\overline{s}$ make impersonal seller more competitive.

Any $\{N, \overline{s}\}$ between N_1 and N_3 but is on the left of N_4 implies that in equilibrium the presence of the impersonal sector hurts personal sellers. Surprisingly, we cannot rule out the cases where personal sellers gain more from competing with the impersonal sector than with another personal seller. If the neutral level of social distance is high and the number of sellers is low, personal sellers sell to buyers socially relatively close and the interaction generates mostly positive social surplus. Therefore personal sellers are in advantageous positions when competing with the impersonal sector, but impersonal seller still gains enough to exist in equilibrium.

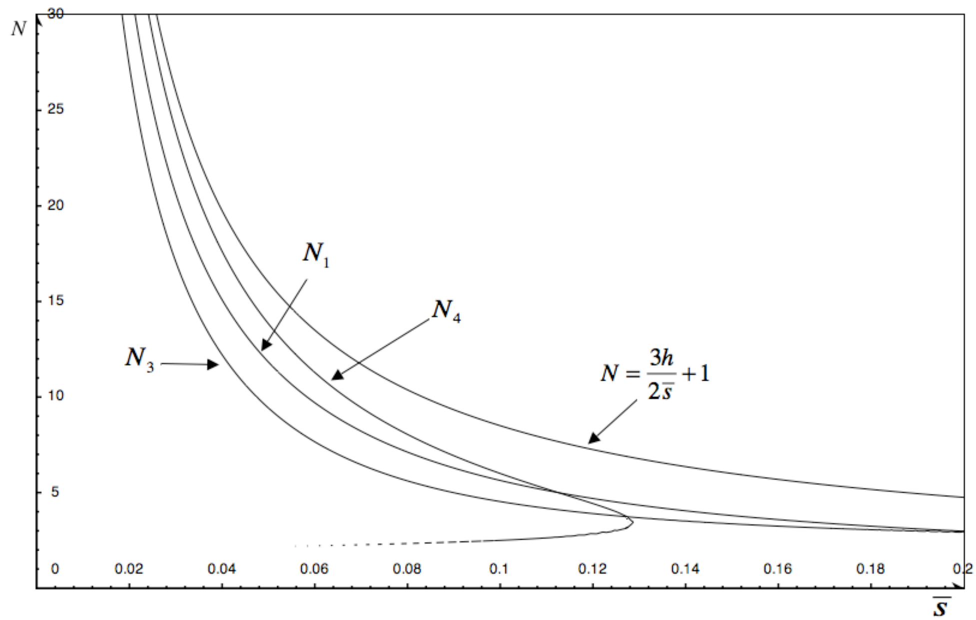


Figure 3.3. Simulated values of N_1 , N_3 and N_4 under parameter values
$\{h = 0.5, k = 1, u = 2, c = 0, l = 10\}$

3.3.2. Buyer's Welfare

Let U_N^b and U_{N-1}^b denote buyers' aggregate utility when there is no impersonal seller and a single impersonal seller respectively.

(3.21)
$$U_N^b = \frac{l}{h}(\sum_{n=1}^{N}(\int_0^{\widehat{s}_{n,N}}(u - \widehat{P}_{n,N} - \overline{k}(a - \overline{s}))da))$$

(3.22)
$$U_{N-1}^b = \frac{l}{h}(\sum_{n=1}^{N-1}\int_0^{\widehat{s}_{n,N-1}}(u - \widehat{P}_{n,N-1} - k_0(a - \overline{s}))da + (u - \widehat{P}_{m,N-1})\widehat{s}_{m,N-1})$$

Increase in $\overline{s}$ benefits buyers as a whole because there will be more positive social surplus from interpersonal interaction. Higher $\overline{s}$ also leads to lower price by decreasing the indifferent marginal buyers' cost of purchasing from the closest sellers.

Not surprisingly, large N brings more competition hence beneficial for buyers.

Let N^b be the solution to $U_N^b = U_{N-1}^b$. Figure 3.4 illustrates the value of N^b against different values of $\overline{s}$. Given $N \le \frac{3h}{2\overline{s}} + 1$, any $\{N, \overline{s}\}$ located on the left of N^b implies $U_N^b < U_{N-1}^b$, and vise versa. This result is intuitive because according to figure 3.3 low $\overline{s}$ and N also make personal sellers worse off when competing with an impersonal seller. Since the game is to some extent zero-sum between buyers and sellers, it is not surprising to see improvement on buyers' welfare under these parameter values. [68]

Recall that the utility of buyer with social distance s units away from the closest seller is $(u - P_n) - \overline{k}(s - \overline{s})$. It is linear in s with slope $-\overline{k}$ and intercept $u - P_n + \overline{k}\overline{s}$. Social surplus declines as social distance increases in the same rate under both cases, but when there is an impersonal sector buyers who purchase from the impersonal seller do not incur further decrease in utility as social distance increases. This result is consistent with Johnson, McMillan, and Woodruff's (2002) observation that impersonal exchange enforced by formal legal system is associated with exchanges of low surplus while more complex exchanges with high surplus are supported by relational contracts.

---

[68] For $N = 2$, it is always true that $U_N^b > U_{N-1}^b$ because if there is an impersonal seller, the single personal seller has no competition in the dimension of social surplus thus enjoys monopolistic power.

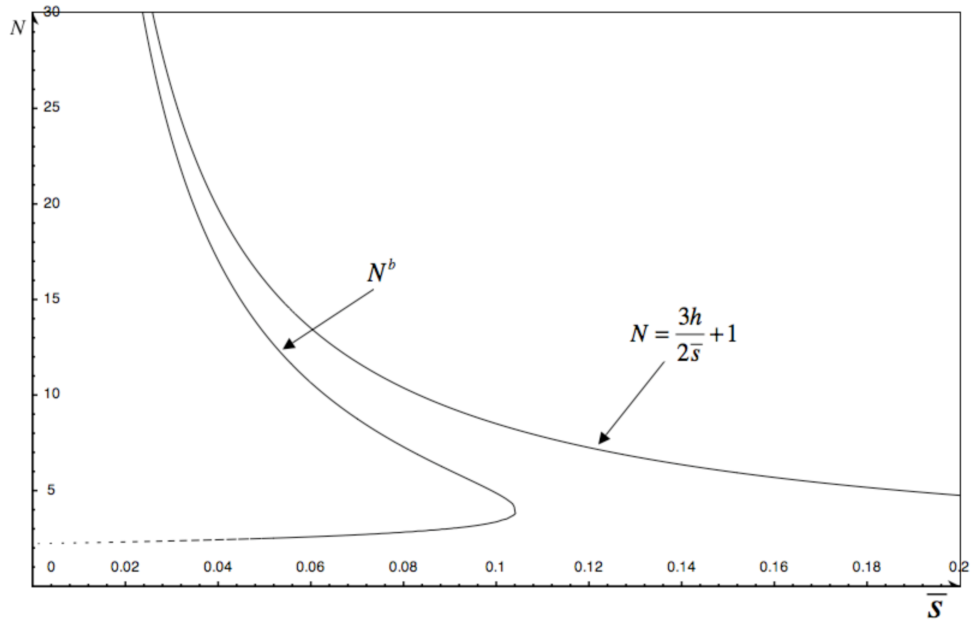


Figure 3.4. Simulated values of N^b under parameter values $\{h = 0.5, k = 1, u = 2, c = 0, l = 10\}$

3.3.2. Inequality

One important aspect that is under-emphasized in the literature is the issue of inequality. With the presence of social surplus that depends on social distance it is especially interesting to investigate the degree of inequality in equilibrium. Because of symmetry I focus on measuring inequality among buyers with social distance from the closest seller ranging from 0 to $\frac{h}{N_p}$.

My analysis follows the spirit of the Gini coefficient (Gini 1912) to construct inequality measure based on welfare.[69] In my model, buyers' utility decreases

[69] While the original Gini coefficient is defined over income that is comparable and additive across individuals, there is issue associated with the validity of aggregating utility in reality because of the cardinal nature of the utility representation. In the model where each buyer is assumed to be identical

monotonically with social distance, so I can calculate the proportion of cumulative utility relative to the total buyers' utility starting from the buyer located farthest (i.e. $s = \frac{h}{N_p}$) and moving inward to the buyer who is closest to the given personal seller. Based on this proportion (the proxy of Lorenz curve) the inequality measure can be calculated in the same way as Gini coefficient. Denote the constructed inequality measure when there is no impersonal sector and one impersonal seller by g_N and g_{N-1}, respectively. Appendix 3.7 gives the details of the construction.

A single factor contributes to inequality: the existence of social surplus. We can see that the constructed measure of inequality is decreasing in economic surplus $u-c$ and $\overline{s}$ but increasing in price because these factors affect each buyer's utility to the same degree, regardless of social distance. On the contrary, the higher $\overline{k}$ is, the higher the inequality, because the degree of social interaction magnifies differences based on social distance. Among buyers who are engaged in impersonal exchange there is complete equality. So the higher the market share of the impersonal seller, the lower the inequality is. As the number of sellers N becomes large, or as $\overline{s}$ increases relative to h, competition drives down price. Consequently, inequality among buyers when there is no impersonal sector decreases unambiguously. However, such increase in N and $\overline{s}$ lower the market share of the impersonal seller, contributing to inequality in the case of a single impersonal seller.

Let N^g be the solution to $g_N = g_{N-1}$. Figure 3.5 shows how the simulated value N^g. Any point $\{N, \overline{s}\}$ located on the left of the curve representing N^g implies

---

ex-ante except for social position, this is less of a concern.

$g_N > g_{N-1}$, and vise versa. In general, when each seller incurs large negative social surplus from interacting with the marginal buyer, relying on impersonal exchange results in more equal outcome among buyers.

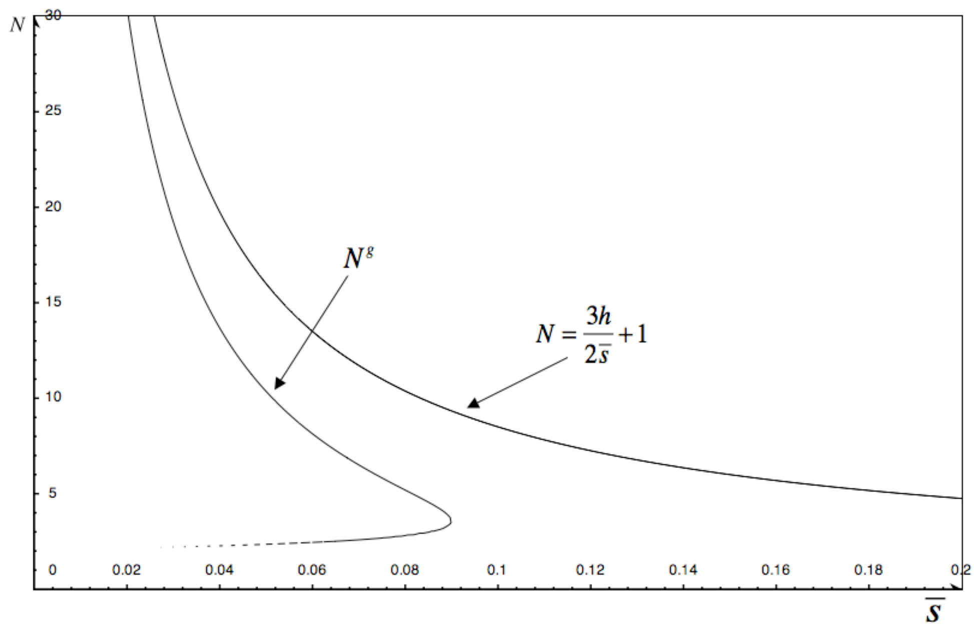


Figure 3.5. Simulated values of N^g **under parameter values** $\{h = 0.5, k = 1, u = 2, c = 0, l = 10\}$

It is informative to combine results from the equilibrium selection and welfare analysis to have a comprehensive understanding of the model's implications. Figure 3.6 graphs N_1, N_3, N^b, and N^g together. To summarize, any point on the left of curve representing N_1 and on the right of curve representing N_3 generates one impersonal seller in equilibrium. Compared to the case of no impersonal sector ($N_p = N$), the presence of an impersonal seller ($N_p = N - 1$) improves buyers' aggregate welfare (points on the left of curve representing N^b), and reduces inequality among buyers (points on the left of curve representing N^g) in many cases. However, there also exist rare cases where the opposite tendencies hold. In particular, in the cases of high $\overline{s}$ and low N, or low $\overline{s}$ and high N (not shown in figure 3.6), compared to having another personal seller, the presence of an impersonal seller in equilibrium makes other personal sellers better off; at the same time buyers' welfare and the degree of equality deteriorate.

These surprising results come from the fact that the presence of impersonal seller permits personal sellers to charge higher price. This is the case when the farthest buyer each personal seller sells to generates social surplus that is neither too high nor too low, so that impersonal seller can attract some buyers to make a high enough profit, but at the same time personal sellers still have some advantages over impersonal seller but not over another personal seller. These cases are more likely in transitional economy where formal market barely starts but personal relationships are still deeply rooted in the society. Similarly, if a society is more homogenous, $\overline{s}$ will be large relative to h. Such society may be located on the right of N_1 but still on the left of curve N^b such that although having impersonal exchange will improve welfare of buyers, the impersonal market exchange will not emerge in equilibrium.

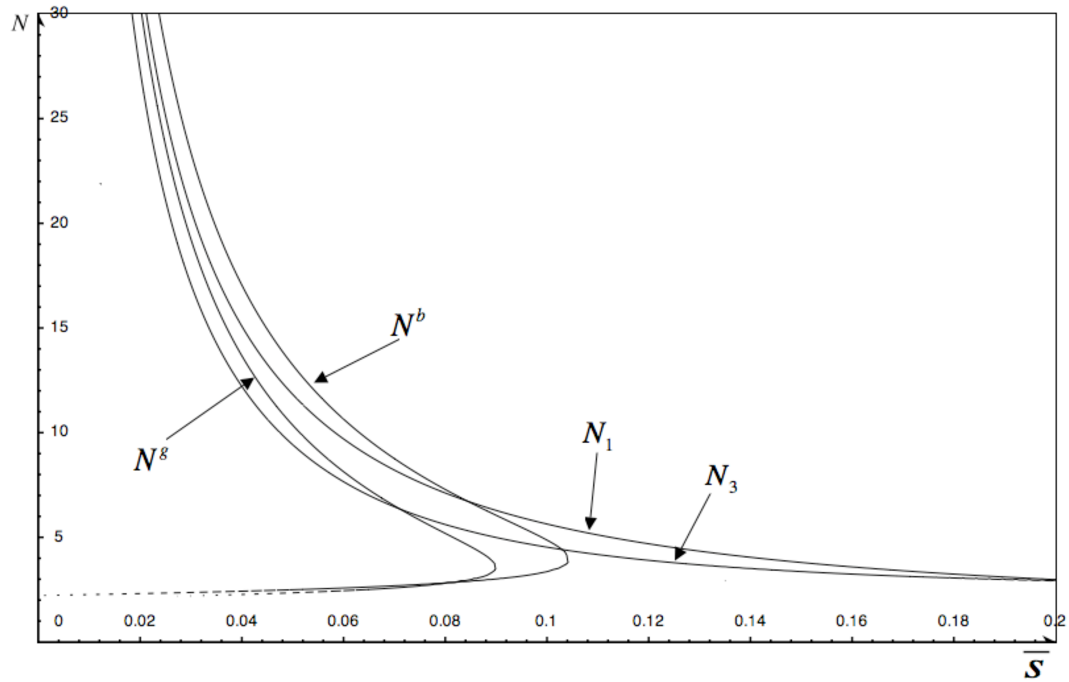


Figure 3.6. Simulated values of N_1, N_3, N^b, and N^g under parameter values $\{h = 0.5, k = 1, u = 2, c = 0, l = 10\}$

3.4. Conclusion

This paper compares personal and impersonal exchange from the perspective of social distance and social surplus. Instead of directly looking at a contract enforcement problem as the literature, I make a reduced-form assumption by introducing social surplus that is specific to personal relationships but not to the impersonal exchange. This social component of the surplus is inversely related social distance, reflecting both different equilibrium payoffs generated by different degrees of altruism, reciprocity or trust in the presence of social dilemma game and the intrinsic matching quality. Different from the main result in the literature that only one of the two forms of exchange can exist in a given society, this paper generates cases where personal relationships and impersonal exchange coexist in equilibrium. Compared to the economy with only personal sellers, the one with an impersonal seller leads to higher welfare and lower inequality among buyers in many cases. But the opposite is true when the impersonal seller barely survives the competition with personal sellers. In general, impersonal exchange is advantageous when the farthest buyer each seller sells to is socially distant. This is the case when there are only a few sellers, the neutral level of social distance is low, or the society is more heterogeneous in general.

The model can be extended along several interesting directions. First, the neutral level of social distance $\overline{s}$ is taken as given in the current model. But in reality it may be endogenously determined by exchange history, market size and the cost of legal enforcement. Similarly, if we assume fixed costs and impose zero profit condition in equilibrium, the number of sellers N is also easily endogenized. Investigating the determinants of $\overline{s}$ and N will help restrict the range of possible equilibria and make sharper predictions. Second, the current model assumes that buyers are uniformly distributed on the circle of social characteristics. This uniform social structure is a simple generalization. In reality, the society may be dominated by a majority group with similar social characteristics so that the distribution of buyers follows a truncated normal distribution. It could also be that there exist several clusters of socially close buyers. A non-uniform distribution of buyers makes personal sellers' location choice nontrivial. It can potentially relate different social structures to whether impersonal exchange will emerge in equilibrium. Third, in reality social distance is never fixed. Although some of the social characteristics are endowed, such as race and gender, others are the results of strategic decision. In particular, Johnson, McMillan, and Woodruff (2002) and Marmaros, and Sacerdote (2006) report evidence that repeated interaction produces closer social distance. Therefore in a dynamic setting the decision to exchange with a particular seller has heavy strategic implications. The possibility of investing in personal relationships is of both empirical relevance and theoretical interest.

Appendix 3.1. Proof of proposition 3.1

For analytical convince, I treat the degree of interpersonal utility k_0 as continuous in the proof. The marginal change of the seller's utility in response to the change in the degree of social surplus is given by

$$\frac{\partial U_0^s}{\partial k_0}\Big|_{s_0=\min(\hat{s}_0,h)} = \frac{l}{h}\left(-\frac{3}{2}(\min(\hat{s}_0,h))^2 + 2\bar{s}\min(\hat{s}_0,h)\right) \tag{3.23}$$

We also know that

$$\frac{\partial \hat{s}_0}{\partial k_0} = -\frac{u-c}{3k_0^2} \tag{3.24}$$

• When $u-c<0$, it is easy to see that $\hat{s}_0 < \frac{2}{3}\bar{s}$ and $\frac{\partial U_0^s}{\partial k_0}\Big|_{s_0=\min(\hat{s}_0,h)} > 0$ for all k_0. The optimal choice is then $k_0 = \bar{k}$.

• When $u-c \geq 0$, we have $\hat{s}_0 \geq \frac{2}{3}\bar{s}$.

· If $\bar{s} \leq \frac{3}{4}h$, it is possible that $\hat{s}_0 \leq \frac{4}{3}\bar{s}$, so we have

$$\frac{\partial U_0^s}{\partial k_0}\Big|_{s_0=\min(\hat{s}_0,h)} \begin{cases} \geq 0, \ if \ k_0 \geq \dfrac{u-c}{2\bar{s}} (equivalently \ \dfrac{2}{3}\bar{s} \leq \hat{s}_0 \leq \dfrac{4}{3}\bar{s}) \\ \leq 0, \ if \ 0 \leq k_0 < \dfrac{u-c}{2\bar{s}} (equivalently \ \hat{s}_0 > \dfrac{4}{3}\bar{s}) \end{cases} \tag{3.25}$$

The seller's utility first decreases and then increases as k_0 goes from 0 to $\bar{k}$. It is therefore not optimal to choose any interior value $0 < k_0 < \bar{k}$. Whether the optimal solution lies on the left or the right boundary depends on the comparison between $U_0^s\Big|_{s_0=\min(\hat{s}_0,h),k_0=0}$ and $U_0^s\Big|_{s_0=\min(\hat{s}_0,h),k_0=\bar{k}}$. Such comparison leads to the conclusion that as long as $\bar{k} \leq k_s = \frac{3(u-c)}{4\bar{s}^2}\left(h - \frac{2}{3}\bar{s} + \sqrt{h^2 - \frac{4}{3}\bar{s}h}\right)$, it is optimal to have $k_0 = 0$; otherwise $k_0 = \bar{k}$.

· If $\bar{s} > \frac{3}{4}h$, we know that $\hat{s}_0 > \frac{4}{3}\bar{s}$. It follows directly that $\frac{\partial U_0^s}{\partial k_0}\Big|_{s_0=\min(\hat{s}_0,h)} > 0$ for all k_0. Thus the optimal degree of interpersonal interaction is $\bar{k}$.

Appendix 3.2. Proof of Lemma 3.1-3.3

Lemma 3.1.

Proof. Personal seller n may face competition from different seller types from two directions and price P_n must respond to different demand curves simultaneously. But by the simple fact that seller n charges uniform price on both sides, the decision on s_n is still symmetric.

Lemma 3.2.

Proof. Assume that personal seller n directly competes with personal seller $n+b$ for any integer $b>1$ on the right hand side, with personal seller $n+1$ located between them. By definition there must exist a marginal buyer who is indifferent between purchasing from personal seller n and seller $n+b$ but strictly prefers these two sellers to personal seller $n+1$. Any buyer located between personal seller n and the marginal buyer strictly prefers to purchase from personal seller n, and those located between the marginal buyer and seller $n+b$ strictly prefer to purchase from seller $n+b$. Hence $U^s_{n+1}=0$. Other things equal, personal seller $n+1$ can do strictly better by moving arbitrarily close to personal seller n and set $P^{'}_{n+1}=P_n$. Therefore there exists a profitable deviation for personal seller $n+1$ that makes seller n directly compete with personal seller $n+1$.

Lemma 3.3.

Proof. If a personal seller n does not compete with impersonal sellers, then she either competes with other personal sellers, or maintains a monopolistic power in her neighborhood. In the first case, there exist marginal buyers who are indifferent between purchasing from seller n and her neighboring personal sellers but strictly prefer these personal sellers than any impersonal seller. Due to symmetry, impersonal sellers do not sell to any buyers. Therefore $U^s_m=0$, contradicting the assumption that $U^s_m>0$. In the second case, there exist marginal buyers who are indifferent between purchasing from seller n and not purchasing any goods. But this case is impossible because $U^s_m>0$ implies that some buyers purchase from impersonal sellers, and by the neutrality of impersonal exchange, the marginal buyers should also strictly prefer purchasing from impersonal sellers.

Appendix 3.3. Proof of Proposition 3.2 and 3.3

When $N_p = N$, it is easy to see that the optimal interior solution of s_n is

(3.26) $$\widehat{s}_n = \frac{B-c}{3\bar{k}} + \frac{2}{3}\bar{s}.$$

Substitute equation (3.26) into the capacity constraint $\sum_{n=1}^{N} \widehat{s}_n = h$ we have

(3.27) $$B = \bar{k}\left(\frac{3h}{N} - 2\bar{s}\right) + c.$$

Given (3.26) and (3.27) we can get the optimal quantity and the corresponding price given in proposition 3.2.

The case when $N_p = N$ is reduced to a standard monopoly's problem and the solution is straightforward.

When $0 < N_p < N$, we can also substitute equation (3.26) to capacity constraint (3.8) and get an intermediate expression for $P_{m,N_p} = B$. Maximizing equation (3.7) given this expression can lead us to (3.13). The rest of the calculations are straightforward.

Appendix 3.4. Proof of Proposition 3.4

Which outcome is equilibrium depends on whether it is strictly better off to become the first impersonal seller, i.e., whether $A = \bar{k}\frac{3lh}{2N^2} - \bar{k}\frac{l}{12h(N-1)}(3h - 2(N-1)\bar{s})^2 > 0$ or not.

Treating N as if it were continuous, we get

$$\frac{\partial(A)}{\partial N} = \bar{k}\frac{l}{2h}(6h^2(\frac{1}{N^3} - \frac{1}{4(N-1)^2}) + \frac{2}{3}\bar{s}^2) \tag{3.28}$$

Given $\bar{s} > \frac{3}{20}h$, we know that the range for the impersonal seller to make positive profit is $N < \frac{3h}{2\bar{s}} + 1 < 11$. In this range $\frac{1}{N^3} - \frac{1}{4(N-1)^2} < 0$ and it is strictly increasing. Depending on the value of $\bar{s}$ relative to h, the utility difference is either monotonic (increasing or decreasing) in N or first decreasing and then increasing in N. Given that when $N \geq \frac{3h}{2\bar{s}} + 1$, impersonal seller has zero utility but personal seller always has strictly positive, in this range $A < 0$. As a result, when $N \neq 2$, there exists a unique point N_1 satisfying $A = 0$. Consequently, if $N < N_1$ it is optimal to have one impersonal seller, otherwise none. When $N = 2$, we can show that a sufficient condition for $A > 0$ is $\bar{s} < \frac{3}{2}(1 - \frac{1}{\sqrt{2}})h$.

Appendix 3.5. Proof of Proposition 3.5

Recall that the necessary and sufficient conditions (3.18) and (3.19) for the equilibrium value of N_p. When $N_p = 0$ only (3.19) is relevant; when $N_p = N$ only (3.18) is relevant. It is easy to apply (3.18) and (3.19) to determine whether these corner values are in equilibrium strategy hence I do not look at them separately.

When $0 < N_p < N$, if we list the left hand side of equation (3.18) and (3.19) according to the order of $N_p = 1,2,...,N-1$, it is easy to see that (3.18) is the lagged term of (3.19) with the opposite sign. For them to be positive simultaneously, it requires that the value of the left hand side of (3.18) goes from positive to negative as N_p increases, and the last term before the sign turns will be the equilibrium strategy.

The first term in (3.18) U_{n,N_p} is decreasing in N_p. The second term U_{m,N_p}, however, has an ambiguous relationship with N_p. Taking derivative of U_{m,N_p} with respect to N_p , we can see that if $N > \frac{6h}{2\overline{s}}$, U_{m,N_p} is decreasing in N_p; if $N \leq \frac{6h}{2\overline{s}}$, it is decreasing in N_p when $N_p < 6h - 2\overline{s}N$ and increasing when $N_p \geq 6h - 2\overline{s}N$.

Given that there exists at least an impersonal seller, the conditions of which have been characterized in proposition 3.4, and the monotonicity described above, we know that the value of $U^s_{n,N_p} - U^s_{m,N_p-1}$ must go from positive to negative once and only once for $0 < N_p < N$. Therefore there exists N_2 satisfying $U^s_{n,N_2} - U^s_{m,N_2-1} = 0$ such that the last positive term before the sign turns will be in equilibrium strategy.

Appendix 3.6. Constructing the Inequality Measure

When there is no impersonal sector, the proxy of Lorenz curve is calculated as

$$L_N = \frac{\int_{s_{cu}}^{\widehat{s}_{n,N}} (u - \widehat{P}_{n,N} - \overline{k}(a - \overline{s}))da}{\int_{0}^{\widehat{s}_{n,N}} (u - \widehat{P}_{n,N} - \overline{k}(a - \overline{s}))da}. \tag{3.29}$$

The corresponding inequality measure is

$$g_N = 1 - \frac{2N}{h}\int_0^{\frac{h}{N}} L_N ds_{cu}. \tag{3.30}$$

When there is one impersonal seller,

(3.31)

$$L_{N-1} = \begin{cases} \dfrac{(u - \widehat{P}_{m,N-1})(\dfrac{h}{N-1} - s_{cu})}{\int_0^{\widehat{s}_{n,N-1}} (u - \widehat{P}_{n,N-1} - \overline{k}(a - \overline{s}))da + (u - \widehat{P}_{m,N-1})\dfrac{\widehat{s}_{m,N-1}}{(N-1)}}, & if\ \widehat{s}_{n,N-1} \le s_{cu} \le \dfrac{h}{N-1} \\ \dfrac{\int_{s_{cu}}^{\widehat{s}_{n,N-1}} (u - \widehat{P}_{n,N-1} - \overline{k}(a - \overline{s}))da + (u - \widehat{P}_{m,N-1})\dfrac{\widehat{s}_{m,N-1}}{(N-1)}}{\int_0^{\widehat{s}_{n,N-1}} (u - \widehat{P}_{n,N-1} - \overline{k}(a - \overline{s}))da + (u - \widehat{P}_{m,N-1})\dfrac{\widehat{s}_{m,N-1}}{(N-1)}}, & if\ s_{cu} < \widehat{s}_{n,N-1} \end{cases}$$

$$g_{N-1} = 1 - \frac{2(N-1)}{h}\int_0^{\frac{h}{N-1}} L_{N-1} ds_{cu}. \tag{3.32}$$

References

Akerlof, George A. 1982. "Labor Contracts as Partial Gift Exchange" The Quarterly Journal of Economics, 97(4): 543-569.

Akerlof, George A. 1984. "Gift Exchange and Efficiency-Wage Theory: Four Views" The American Economic Review, 74(2, Papers and Proceedings of the Ninety-Sixth Annual Meeting of the American Economic Association): 79-83.

Akerlof, George A. 1997. "Social Distance and Social Decisions" Econometrica, 65(5): 1005-1027.

Akerlof, George A. and Rachel E. Kranton. 2000. "Economics and Identity" The Quarterly Journal of Economics, 115(3): 715-753.

Akerlof, George A. and Janet L. Yellen. 1990. "The Fair Wage-Effort Hypothesis and Unemployment" The Quarterly Journal of Economics, 105(2): 255-283.

Benjamin, Daniel J. 2008. "Social Preferences and the Efficiency of Bilateral Exchange" Cornell University and Institute for Social Research Mimeo.

Bernhard, Helen, Ernst Fehr, and Urs Fischbacher. 2006. "Group Affiliation and Altruistic Norm Enforcement" The American Economic Review, 96(2): 217-221.

Bohneta, Iris and Bruno S. Frey. 1999. "The Sound of Silence in Prisoner's Dilemma and Dictator Games" Journal of Economic Behavior and Organization, 38: 43-57.

Bowles, Samuel and Herbert Gintis. 2004. "Persistent Parochialism: Trust and Exclusion in Ethnic Networks" Journal of Economic Behavior & Organization, 55: 1-23.

Buchan, Nancy R., Eric J. Johnson, and Rachel T. A. Croson. 2006. "Let's Get Personal: An International Examination of the Influence of Communication, Culture and Social Distance on Other regarding Preferences" , 60: 373-398.

Carmichael, H. L. and W. B. MacLeod. 1997. "Gift Giving and the Evolution of Cooperation" International Economic Review, 38(3): 485-509.

Charness, Gary and Uri Gneezy. 2008. "What's in a Name? Anonymity and Social Distance in Dictator and Ultimatum Games" Journal of Economic Behavior and Organization, 68: 29-35.

Charness, Gary, Ernan Haruvy, and Doron Sonsino. 2007. "Social Distance and Reciprocity: An Internet Experiment" Journal of Economic Behavior and Organization, 63: 88-103.

Currie, Janet, Jeffrey Grogger, Gary Burtless, and Robert F. Schoeni. 2001. "Explaining Recent Declines in Food Stamp Program Participation [with Comments]" Brookings-Wharton Papers on Urban Affairs: 203-244.

Dixit, Avinash. 2003. "Trade Expansion and Contract Enforcement" The Journal of Political

Economy, 111(6): 1293-1317.

Gini, Corrado. 1912. "Variabilità e mutabilità" Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi (1955).

Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies" The Journal of Political Economy, 102(5): 912-950.

Hart, Oliver and John Moore. 2008. "Contracts as Reference Points" Quarterly Journal of Economics, 123(1): 1-48.

Hoffman, Elizabeth, Kevin McCabe, and Vernon L. Smith. 1996. "Social Distance and Other-regarding Behavior in Dictator Games" The American Economic Review, 86(3): 653-660.

Kimbrough, Erik, Vernon L. Smith, and Bart J. Wilson. 2008. "Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long-Distance Trade" American Economic Review, 98(3): 1009-1039.

Kranton, Rachel E. 1996a. "Reciprocal Exchange: A Self-Sustaining System" The American Economic Review, 86(4): 830-851.

Kranton, Rachel E. 1996b. "The Formation of Cooperative Relationships" Journal of Law, Economics, & Organization, 12(1): 214-233.

Leeson, Peter T. 2008. "Social Distance and Self-Enforcing Exchange" The Journal of Legal Studies, 37(1): 161-188.

Macaulay, Stewart. 1963. "Non-Contractual Relations in Business: A Preliminary Study" American Sociological Review, 28(1): 55-67.

Marmaros, David and Bruce Sacerdote. 2006. "How do Friendships Form?" The Quaterly Journal of Economics, 121(1): 79-119.

Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma" The American Economic Review, 73(5): 1023-1035.

Salop, Steven C. 1979. "Monopolistic Competition with Outside Goods" The Bell Journal of Economics, 10(1): 141-156.

Simon Johnson, John McMillan, and Christopher Woodruff. 2002. "Courts and Relational Contracts" Journal of Law, Economics, & Organization, 18(1): 221-277.

Sobel, Joel. 2002. "Can we Trust Social Capital?" Journal of Economic Literature, 40(1): 139-154.

Sobel, Joel. 2006. "For Better Or Forever: Formal Versus Informal Enforcement" Journal of Labor Economics, 24(2): 271-297.

Stuber, Jennifer and Karl Kronebusch. 2004. "Stigma and Other Determinants of Participation in TANF and Medicaid" Journal of Policy Analysis and Management, 23(3): 509-530.