INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Durable Goods Consumption:  
From Micro Foundation  
To Macro Dynamics  

DISSERTATION

Presented in Partial Fulfillment of the Requirements for  
the Degree Doctor of Philosophy in the  
Graduate School of The Ohio State University  

By  
Xiaojun Wang, B.A., M.A.  

*******  
The Ohio State University  
2001

Dissertation Committee  
Paul D. Evans, Adviser  
Pok-Sang Lam  
Masao Ogaki  

Approved by  
Paul D. Evans  

Adviser  
Department of Economics
Rational households adopt (S,s)-type rule for durable goods consumption due to the presence of non-convex adjustment costs. In this thesis, I numerically investigate the micro theoretical model, empirically test (using two methods) the model with micro data, and then analyze the implication of micro inertial behavior on the macro dynamics.

First, A set of decision rules is numerically solved using the continuous time optimal portfolio choice model developed in Grossman and Laroque (1990) *Econometrica* paper. The key implication of model is that households facing higher adjustment costs adopt wider (S,s) rules.

Second, a micro panel data set on the value of households' automobile stocks is used to estimate the (S,s) decision rules for durable goods purchases and sales. The data is first analyzed with a threshold auto-regression model. Preliminary evidence is found in support of the theoretical model. Then a much more elaborate model, i.e., mixture distribution maximum likelihood model is introduced. In this new
model, households monitor the wealth-durables ratios, (the logarithm of) which are assumed to follow controlled diffusion processes. Exploiting the resulting ergodic distribution, I estimate the relation between the boundary conditions and households’ demographic characteristics with a mixture distribution model. I find strong evidence of heterogeneity in households’ \((S,s)\) rules. The observed heterogeneity is consistent with the theory. Moreover, the ergodic distribution implied by the estimated model fits the sample distribution well and much better than those implied by models that ignore heterogeneity.

Finally, based on the parameters estimated in the mixture distribution model, I simulate the aggregate durable goods expenditure dynamics after random shocks. This exercise shows that the aggregate of individual \((S,s)\) rules is capable of producing macro dynamics similar to those observed in the data.

This thesis is the first in the literature to use the mixture distribution method to recover the household’s \((S,s)\) rules. In addition, findings in this thesis show the importance of taking care of aggregation problems while testing a micro model with macro data.
This is Dedicated to My Parents
ACKNOWLEDGMENTS

First of all, I am deeply grateful for the general guidance of my committee members. My adviser, Dr. Paul D. Evans, has not only provided excellent academic advice, but also enormous emotional support during the ups and downs of this project. He shares both his economics knowledge and lifetime philosophy with me, which I enjoy very much and shall benefit from for the rest of my life. Without him, this work will not exist. It is a privilege for me to have him as my academic adviser.

I also highly appreciate continuous support by the other two committee members, Dr. Pok-Sang Lam and Dr. Masao Ogaki. Dr. Lam’s early work is one of the motivations of this project. Several long discussions with him help me clarify a lot of important issues in the subject. Dr. Ogaki suggested the overhaul of the modeling method in my proposal, which greatly improves the academic quality of this work. I can not overstate my appreciation for their warm-hearted support.

My thanks also goes to Dr. Stephen Cosslett, Dr. Eric Fisher, Dr. Belton Fleisher, Dr. J. Huston McCulloch and Mr. Yong Yu. I appreciate their helpful
comments and suggestions during various stages of this project. I also thank the Economics Department staff members, especially Chris Colflesh, Jo Ducey and John-David Slaughter for their help throughout these years. Financial support from the Economics Department as well as the Graduate School is also greatly acknowledged.

I would like to thank my fellow graduate students in both the Economics and Agricultural Economics Departments for making my life in this foreign country less stressful and more enjoyable. They are: Chi-Young Choi, Wen-Hua Hsieh, Suzie Greenhalgh, Justin Pestrue, Ken Brevoort and Thomas Chapman. I wish you all the best in your life.

Finally, I would like to thank my family here and far away. My wife, Qing Liu, has supported and helped me in each and every aspect during these tough years. My parents, brother, and sister in China, have loved me unconditionally for the past 31 years. I cannot be who I am today without them. I want them to share my joy of success.
VITA

January 21st, 1970 ................................. Born-Hubei, P.R.China

1992 ...............................B.A. Economics, Wuhan University, P.R.China

1995 ..............................M.A. Economics, Wuhan University, P.R.China

1996 ..............................M.A. Economics, The Ohio State University

1995-present ...........Graduate Teaching Associate, The Ohio State University

FIELDS OF STUDY

Major Field: Economics

Studies in:

1 Macroeconomics

2 Econometrics

3 China Labor Economics
# Contents

Abstract i

Dedication iii

Acknowledgments iv

Vita vi

List of Tables x

List of Figures xi

1 INTRODUCTION 1

2 THEORETICAL MODEL AND NUMERICAL SOLUTIONS 8

2.1 Theoretical Model 8

2.2 Numerical Simulations 14

2.3 Ergodic Distribution 21
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Numerical Simulation Results</td>
<td>83</td>
</tr>
<tr>
<td>2 Sample Statistics</td>
<td>84</td>
</tr>
<tr>
<td>3 Threshold Estimation Result</td>
<td>84</td>
</tr>
<tr>
<td>4 Determinants of Bandwidth</td>
<td>85</td>
</tr>
<tr>
<td>5 Minimum Requirement on Thresholds</td>
<td>85</td>
</tr>
<tr>
<td>6 Direct Estimation with Observed Heterogeneity</td>
<td>86</td>
</tr>
<tr>
<td>7 Direct Estimation without Observed Heterogeneity</td>
<td>87</td>
</tr>
<tr>
<td>8 Homogeneous Model</td>
<td>87</td>
</tr>
<tr>
<td>9 Maximum Likelihood Estimation Results Heterogeneous Model</td>
<td>88</td>
</tr>
<tr>
<td>10 Model without Observed Heterogeneity</td>
<td>89</td>
</tr>
<tr>
<td>11 Predicted (S,s) Rules</td>
<td>89</td>
</tr>
<tr>
<td>12 Percentage of Population Who Adjust in Steady State</td>
<td>89</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Ergodic Distribution</td>
<td>90</td>
</tr>
<tr>
<td>2 Empirical Density</td>
<td>91</td>
</tr>
<tr>
<td>3 Sample Path of Optimal (S,s) Behavior</td>
<td>92</td>
</tr>
<tr>
<td>4 Ergodic Distribution and $\gamma$</td>
<td>93</td>
</tr>
<tr>
<td>5 Sample Histograms</td>
<td>94</td>
</tr>
<tr>
<td>6 Goodness of Fit Comparison</td>
<td>95</td>
</tr>
<tr>
<td>7 RDPIPC, RDEPC(M), and Recessions</td>
<td>96</td>
</tr>
<tr>
<td>8 Simulated Impulse Responses</td>
<td>97</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

Households' holdings of durable consumption goods are usually modeled as following an \((S,s)\)-type rule because the purchases and sales of such goods entail adjustment cost. These \((S,s)\) rules depend not only on the income dynamics of the households but also on their demographical characteristics. Most of previous work in this field has exploited data on households' purchase and sale behavior. By contrast, I make use of results from the latest literature and estimate the model with data on the stock of a durable good. It represents the first direct attempt to test the model proposed in Grossman and Laroque (1990).

It has been a long recognized fact that changing durable goods stock incurs adjustment cost. When this cost is a lump-sum fixed amount, or a non-convex function of the size of the adjustment, households will not adjust even though their current stocks are not at the optimal level. On the contrary, they will wait until
the deviation is large enough and then execute adjustment. Upon the adjustment, durable stock is reset at the optimal level, and the household may stay inactive for quite a while before another adjustment becomes necessary. That is to say, at least implicitly, each household has an inaction range for durable goods stock (which constitutes the so-called (S,s) band), and an optimal level which it adjusts to when a change is called for.\(^1\) Deviation from the optimal level may come from several sources. Income shock and durable depreciation are the main driving forces. Demographical changes such as getting married or expecting a newborn also prompt durable goods adjustment. One of the major goals of this paper is to shed light on how these changes affect households' decision rule.

Recovering the households (S,s) decision rules on durable goods consumption has at least two purposes. First, it enhances our understanding of the mechanism by which micro individuals make consumption decisions. For example, how a reduction in adjustment cost may encourage people to buy more often. Second, individual inertial behavior turns out to have strong implications on the aggregate dynamics. Since each household usually is only active in the durable goods market once in several years, the aggregate expenditure may exhibit dynamics completely different

---

\(^1\) The optimal level, or internal return point, or target level, all of which will be used interchangeably throughout the paper, is usually different from the optimal level without adjustment cost. See Bertola and Caballero (1990).
from the individual counterpart. As a result, it might be inappropriate to test micro models with macro data without considering the possible aggregation problem. I will attempt to tackle both issues in such order.

This paper is the first attempt to estimate the \((S,s)\) rule with stock values instead of flow values. That is to say, I use durable stock data rather than commonly used expenditure data. For example, in the second method proposed below, i.e., the mixture distribution estimation, this is achieved by exploiting the ergodic distribution of durable stock to which each individual converges to. If there is no aggregate shock and households make independent decisions, this guarantees a stable cross-sectional distribution of durable goods stock as the sample size, in either a cross section or time series, approaches infinity.\(^2\) As a result, each observation on the stock holding can be viewed as a random draw from the household’s ergodic distribution, and it is a sufficient statistic to estimate the cross-sectional distribution of the \((S,s)\) rules. In addition, I also consider the effect of households’ demographic characteristics on the parameters of the ergodic distribution, i.e., targets and band widths. The only other research that does this, as far as I know, is Attanasio (1999). There are two major difference between his work and mine. First, his model requires expenditures data, and exact time and amount of each purchase or resale (adjustment). All I use is durable stock data, and no information on the exact adjustment is needed. Second,

\(^{2}\) This is an application of the Glivenko-Cantelli Theorem. See Billingsley (1986). This result also holds when households \((S,s)\) decision rules are drawn iid from a non-degenerated distribution.
his estimation method is very similar to Lam (1991), using a Tobit model to infer on boundary conditions. In this paper, instead, I construct a mixture distribution model and estimate the boundary conditions with Maximum Likelihood.

The methodology adopted here has several advantages over those used in previous studies. First, it is a more efficient. Eberly (1994) use the observed maximum and minimum values to be the boundaries for the \((S,s)\) rule. This certainly ignores information carried in those observations between the extremes. In the current framework, every observation is used to extract information on the ergodic distribution, and they are treated equally. Second, I assume that the true \((S,s)\) rule is unobservable and the observed values only provide an upper limit or lower limit. This method is more robust to measurement errors and usually performs better than that assumes observable bounds (See Attanasio (1999).). Finally, this method is very flexible with data. It uses stock values instead of flow values, which usually requires detailed information on the individual purchases record. It can also take on rotating samples (such as Consumer Expenditure Survey, CEX) and non-continuous samples, even panel data with lots of missing observations. This makes it possible to be used in a much wider field.

The motivation of this work comes from two lines of literature. One derives the \((S,s)\) rule as the optimal method of holding stocks of consumer durable goods. Grossman and Laroque (1990) and Bar-Ilan and Blinder (1988) provide a theoretical foundation for the \((S,s)\) rule when there are transaction costs of adjusting consumer
durable goods stock. Lam (1991) estimates an (S,s) model in levels with panel data. However, he only considers unobserved heterogeneity and ignores the effect of households' demographic characteristics (observed heterogeneity). Attanasio (1999) takes this effect into account, but he uses the durable goods - non-durable goods ratio as the state variable. Both of them use durable expenditure data to infer on the boundary conditions.

The second line also starts from Grossman and Laroque (1990), but is further developed in Eberly (1994). It can be shown that if a state variable follows a controlled diffusion process, it has a long-run, or ergodic, distribution. Moreover, if the population is homogeneous, and we observe enough individuals, the cross-sectional distribution of durable goods holdings is the ergodic distribution. Eberly uses this feature to evaluate the macroeconomic features of the model and concludes that a homogeneous model provides a good fit to her sample.

The main goal of this work is two-folded. On the micro side, I will investigate how strong the heterogeneity is in individual decision making, and how those observable heterogeneity, adjustment costs in particular, affect the decisions. On the macro side, I will show how individual inertial behavior renders volatility and persistence in aggregate dynamics.

---

3 See previous footnote 2. The same thing holds for one individual observed repeatedly over a long enough time period.
The optimality of (S,s) rule in the presence of adjustment costs is taken for granted for the most part of this work, after all, there is no need to reinvent the wheel. One can find very good expositions on that issue in Grossman and Laroque (1990) (in a stochastic framework) and Bar-Ilan and Blinder (1988) (in a deterministic framework). Rather, it is the starting point of my analysis.

Major findings of this work are summarized as follows:

- There are significant amount of heterogeneity, both observed and unobserved, in households’ (S,s) choices. This contrasts the acceptance of homogeneity in Eberly (1994).

- Observed demographic variables affect households’ (S,s) choices in a way consistent with the theory. In particular, higher adjustment cost does prompt people to increase inaction range.

- The estimated heterogeneous model fits the sample much better than the homogeneous counterpart.

- Simulation based on the estimated heterogeneous model exhibits aggregate dynamics that conform well with the observed macroeconomic data. In particular, volatility and persistence in aggregate expenditures is mainly due to changes in the number of adjustment, not average expenditure.
The rest of the thesis is organized as follows. The next chapter reiterate the theoretical model *a la* Grossman and Laroque (1990) and optimal decision rules are numerically solved. As a preliminary investigation, a threshold auto-regression model is presented in chapter 3. Chapter 4 presents the mixture distribution model, which is the key contribution of this work. Chapter 5 reports the estimation results from chapter 4. Chapter 6 simulates aggregate expenditure on durable goods based on the estimated model of chapter 4. This simulated dynamics is then analyzed and compared with the observed aggregate data. Chapter 7 concludes and outlines future work. All referred tables and figures are gathered at the back of the paper.
Chapter 2

THEORETICAL MODEL AND NUMERICAL SOLUTIONS

In this chapter, I first set up the theoretical framework, which is in short an optimal stopping problem with controlled diffusion process. Then, numerical simulations are carried out to illuminate relations between (S,s)-type behavior and underlying parameters. Finally, an ergodic distribution is derived to show how each individual behaves in the long run, and also its implication on the cross-section distribution.

2.1 Theoretical Model

The basic setup of the model is the same as in Grossman and Laroque (1990), to which the reader is referred to for additional motivation and clarification. The
following outlines important structures that will be used in this chapter. Rigorous proofs can be found in the original work and its early version (both are cited in the reference).

An infinitely lived household derives utility from possessing a composite durable good $K$. At the beginning of its lifetime, the household is endowed with certain amount of financial assets as well as some initial level of durable stock $K$. The financial assets include risk free asset $A$ and/or risky asset $X$. Risk free asset has a constant return rate of $r_f$, and the value of the risky asset evolves as a geometric Brownian motion specified below. Financial market is perfect in that households are free to change their portfolio composition with no transaction charges. However, adjusting durable stock incurs capital loss proportional to the amount to be resold. When no adjustment occurs, $K$ depreciates exponentially at a rate of $\delta$. To summarize the budget constraints:

$$ W_t = K_t + A_t + X_t $$

(2.1)

where $W_t$ is total wealth as of time $t$. When the household doesn’t adjust its durable stock, $W_t$ evolves as:

$$ dW_t = -\delta K_t dt + r_f A_t dt + X_t (\hat{\mu} dt + \sigma d\omega_t) $$

(2.2)

where $\hat{\mu}$ is the instantaneous drift and $\sigma$ the instantaneous standard deviation. $\omega_t$ is a Brownian motion without drift. If the household decides to adjust its durable
stock, then:

\[ W_{\tau} = W_{\tau^{-}} - \lambda K_{\tau^{-}} \quad (2.3) \]

where \(\tau\) is the first stopping time and \(\tau^{-}\) is the moment right before \(\tau\). \(\lambda\) is the size of adjustment cost measured as a fixed proportion of the preceding durable stock.

To complete the budget side, I need to impose the no bankruptcy condition:

\[ W_t - \lambda K_t \geq 0 \quad \text{for all } t \quad (2.4) \]

Consider a household maximizing the expected value of a time separable utility function \(E \int_0^\infty e^{-\theta t} U(K_t)dt\) where \(\theta > 0\) is the subjective discount rate. The consumption service flow is taken to be proportional to the durable stock \(K_t\). Given initial conditions \((W_0, K_0)\), the household needs to find the path of \((X_t, K_t)\) and first stopping time \(\tau\) (at which it adjusts the durable stock) to maximize its lifetime utility subject to the budget constraints.

Assuming constant relative risk aversion (CRRA) instantaneous utility function, and let \(V(W, K)\) be the supremum of the expected utility that the household can achieve from the initial conditions \((W, K)\). Bellman equation can be formulated as:

\[ V(W, K) = \sup_{c,\tau, X_t} E[\int_0^\tau e^{-\theta t} \frac{K_t^a}{a} dt + e^{-\theta \tau} V(W_{\tau} - \lambda K_{\tau}, c)] \quad (2.5) \]

where \(\tau\) is the first stopping time from date 0, \(c\) is the value of new durable purchased,
and \( a \) is relative risk aversion coefficient, \( a < 1 \) and \( a \neq 0 \).

This problem can be simplified by using the following transformations:

\[
y = \frac{W}{K} - \lambda \quad (2.6)
\]

\[
x = \frac{X}{K} \quad (2.7)
\]

\[
h(y) = K^{-a}V(W, K) = V(\lambda + y, 1) \quad (2.8)
\]

\[
\bar{\theta} = \theta + a\delta \quad (2.9)
\]

\[
\mu = \mu - r_f \quad (2.10)
\]

where equation (2.8) exploits the homogeneity property of \( V(W, K) \). Let

\[
M = \sup_c \left( \frac{W_c - \lambda K_c}{c} \right) - a h\left( \frac{W_c - \lambda K_c}{c} - \lambda \right) \quad (2.11)
\]

\[= \sup_y (y + \lambda)^{-a} h(y) \quad (2.12)
\]

then the above problem is simplified to:

\[
h(y) = \sup_{y, x(t)} \mathbb{E} \left[ \int_0^t \frac{e^{-\delta t}}{a} \, dt + e^{-\delta t} M y_t^a \right] \quad (2.13)
\]

and subject to:

\[
dy = (x \mu + r(y + \lambda - 1)) \, dt + x \sigma \, d\omega \quad (2.14)
\]

\[^1\text{To ensure the finiteness of the above value function,}
\]

\[
\beta = \theta - ar_f - \frac{\mu^2}{2\sigma^2} \frac{a}{1-a} > 0
\]
and

\[ y_t \geq 0 \quad \text{at all } t \]  \hspace{1cm} (2.15)

Let M be exogenously fixed satisfying equation (2.12), the solution to the system of equations (2.13), (2.14), and (2.15) is:

(i) If \( h(y) > My^a \), then it is optimal not to stop (i.e., \( r \neq 0 \)), \( h(y) \) is twice continuously differentiable except possibly at \( y = 1 - \lambda \) and:

\[
\sup_x \left[ \frac{h''(y)}{2} Var(dy) + h'(y)E(dy) - \delta h(y) + \frac{1}{a} \right] = 0
\]  \hspace{1cm} (2.16)

where \( Var(dy) = x^2 \sigma^2 \) and \( E(dy) = r(y + \lambda - 1) + x \mu \).

(ii) If \( h(y) = My^a \), then stop (i.e., \( r = 0 \)) and, in the interior of the set \( \{y|h(y) = My^a\} \),

\[
\sup_x \left[ \frac{h''(y)}{2} Var(dy) + h'(y)E(dy) - \delta h(y) + \frac{1}{a} \right] \leq 0
\]  \hspace{1cm} (2.17)

The above result can be derived by invoking continuity property of state space \( y \) and policy space \( x \), and then apply Ito’s Lemma.

Optimal portfolio choice of \( x \) can be derived from above as:

\[ x(y) = \frac{-h'(y)}{h''(y)} \frac{\mu}{\sigma^2} \]  \hspace{1cm} (2.18)

It should be clear by now that this is a problem of optimal control over a diffusion process. Grossman and Laroque (1990) show that solution to the above system is characterized by three values: lower bound \( y_1 \), upper bound \( y_2 \), and internal return
point $y^*$. The Household monitors the only state variable $y$, which by definition is determined by total wealth and current durable stock. If $y$ ever reaches either bound, it re-sells the current durable stock and purchases a new one (either with higher value in case of an upward adjustment, or with lower value in case of a downward adjustment), and the state variable assumes the value of $y^*$. Substituting equation (2.18) into equation (2.16) and imposing the value matching conditions and smooth pasting conditions at the boundary points, I obtain the following method of solving for $(h(y), y_1, y_2, y^*)$:

$$\frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 \left( \frac{h''(y)}{h'(y)} \right)^2 + \frac{\sigma(\gamma + 1)h'(y)}{h''(y)} - \frac{\sigma}{\sigma + 1} = 0$$  \hspace{1cm} (2.19)

\text{for } y \in [y_1, y_2], \ y \neq 1 - \lambda; \hspace{1cm}

$$h(y) \geq M y^a \text{ for all } y;$$  \hspace{1cm} (2.20)

$$h(y_i) = M y_i^a \text{ for } i = 1, 2;$$  \hspace{1cm} (2.21)

$$h'(y_i) = a M y_i^{a-1} \text{ for } i = 1, 2;$$  \hspace{1cm} (2.22)

and $M$ must satisfy:

$$M = (y^* + \lambda)^{-a} h(y^*) = \sup_y (y + \lambda)^{-a} h(y)$$  \hspace{1cm} (2.23)

and the internal return point can be solved using equation (2.23) too.

---

2 In the general model of Bertola and Caballero (1990), hitting the upper bound returns the state variable to a different internal return point from hitting the lower bound. However, in the specific model of Grossman and Laroque (1990), hitting either bound return the state variable to a common internal return point. This is because adjustment cost only depends on the preceding durable stock in a proportional way, and it does not depend on the size of the adjustment as assumed in Bertola and Caballero.
Equations in (2.21) are the so-called value matching conditions. Intuitively, when the state space and policy space are both continuous, the household must be indifferent between stopping and non-stopping at the boundaries. The left-hand-side is the value of non-stopping, and the right-hand-side is the value of stopping. If these equalities do not hold, the household should advance or postpone its stopping decision. Equations in (2.22) are called smooth pasting conditions. By differentiability of the transformed value function \( h(y) \) and continuity of \( y \) and \( x \), marginal benefit of stopping and non-stopping should also be the same. Otherwise, \( y \) has the tendency to break through the boundaries when it gets infinitesimally close.\(^3\)

Even though no analytic solution can be found for the system (2.19)-(2.23), numerical simulations can be carried out to find the approximate decision rules under different parameters calibration. The exact algorithm is briefly discussed in appendix A, and the simulation results are reported in the next chapter.

### 2.2 Numerical Simulations

In this subsection I simulate the model described in the preceding subsection and see how underlying parameters affect household decision rule. Table 1 shows major results from the simulation for the value of \( \lambda \) equal to 0.01, 0.05 and 0.10 respectively.

\(^3\)See Bertola and Caballero (1990) page 245-247 for more on both conditions.
For each case, constant relative risk aversion coefficients range from 0.9 to 3.1. Other parameters are calibrated as follows (all follow the original paper): durable depreciation rate $a$ is 10 percent per year, risk free rate $r_f$ is 1 percent per year, subjective discount rate $\theta$ is 1 percent per year, mean excess rate of risky asset $\mu$ is 5.9 percent per year, and its standard deviation $\sigma$ is 22 percent per year.

The following columns are included:

- Column 1: $A = 1 - a$, a measure of constant relative risk aversion;

- Column 2: $(y_1, y^*, y_2)$ triple, i.e., lower bound, internal return point and upper bound;\(^5\)

- Column 3: $y_2 - y_1$, a measure of bandwidth;

- Column 4: $E[T(y^*)]$, expected first hitting time evaluated at the internal return point;

- Column 5: $(\frac{K}{W})^*$, durable/wealth ratio evaluated at the internal return point;

- Column 6: $(\frac{A}{W})^*$, risk-free-asset/wealth ratio evaluated at the internal return point;

---

\(^4\)This is deliberately chosen to skip the nuisance case of $A = 1$, in which case the whole model falls apart. Also the range goes up to 3.1, which is higher than what is used in the original paper by Grossman and Laroque. This is just taking account of the latest estimates by Ogaki and Reinhart (1998).

\(^5\)Note that in Grossman and Laroque, they report this column including a $\lambda$; Here I do not.
- Column 7: \((\frac{\chi}{W})^*\), risky-asset/wealth ratio evaluated at the internal return point;

- Column 8: drift in income;

- Column 9: standard deviation in income.

Column 1-3 is self-obvious. Column 4 reports mean first hitting time from the internal return point, i.e., expected length of time between household’s durable purchases. Following Karlin and Taylor (1981, p192), let \(T(y)\) be the length of time it takes to hit a boundary (either \(y_1\) or \(y_2\)) starting from any internal point \(y\). Certainly \(T(y)\) is a random variable. We are only interested in its mean, that is, average time it takes to hit a boundary. Let

\[
V(y) = E[T(y)]
\]  

Then \(V(y)\) satisfies the following differential equation:

\[
-1 = V'(y)E[dy] + \frac{1}{2} V''(y)Var[dy]
\]

with boundary conditions \(V(y_1) = V(y_2) = 0\). This is solved numerically (see appendix A for details) and \(V(y^*)\) is reported in Column 4.

Column 5-7 reports the share of total wealth of durable goods, risk-free asset and risky asset, respectively. They are evaluated right after each purchase, i.e., at \(y^*\). Straightforward algebra shows:

\[
(\frac{K}{W})^* = \frac{K_{\tau}}{W_{\tau} - \lambda K_{\tau}} = \frac{1}{y^*},
\]  

16
\[ \left( \frac{A}{W} \right)^* = y^* + \lambda - 1 - x^* = \frac{y^* + \lambda - 1 + \frac{k'(y^*)}{k''(y^*)} \mu}{y^*}, \]  
\[ \left( \frac{X}{W} \right)^* = -\frac{\lambda + \frac{k'(y^*)}{k''(y^*)} \mu}{y^*} \]  
\[ (2.27) \]  
\[ (2.28) \]

Note all three equations also hold at any \( y \), so we can use them to show the dynamics of portfolio adjustment.

The last 2 columns report the drift and standard deviation per unit time of household’s income flow from holding both the risk-free and risky assets. Here I assume that the only source of income per period is from interest and dividend. The household only faces a re-investment portfolio choice problem when it does not adjust its durable stock. Drift and standard deviation are calculated as a weighted-average of income sources. As mentioned before, these two vary constantly across the state space, but only the value at \( y^* \) is reported.

Result Analysis: First, let’s focus on each individual category of \( \lambda \). It is easy to find the following patterns: a). Upper bound and lower bound are not symmetric, the distance between upper bound and internal return point is usually greater than that between lower bound and internal return point. This is not much of a surprise because there is a positive drift in total wealth. As a result, households expect to hit upper bound more frequent (or alternatively, with higher probability). To correct for this type of asymmetry in wealth drift, it is optimal for households to have more room on the upper side of the threshold. However, they would not go as far as equalizing the probability of hitting either bound because adjusting upward
incurs more real capital loss than adjusting downward. b). Households with higher risk aversion have smaller bandwidth, but average time between purchases is not a monotonic function of risk aversion. The former is because more risk averse households usually hold less portion of their wealth in risky asset, this reduces the drift in wealth, and in turn reduces the speed the state variable $y$ travels in the band. However, smaller band does not necessarily imply more frequent purchase. For example, when $\lambda$ equals to 0.01 and $A$ increases from 0.9 to 1.1, bandwidth reduces from 23.57 to 8.75, but the time between adjustments increases from 3.05 to 3.20. The reason is that even though households with higher risk aversion have smaller bandwidth, their drift of wealth decreases too, these two effects jointly determine how frequent households adjust, and the ultimate effect is ambiguous. Nevertheless, for most of the constant risk aversion coefficients listed in the table, households with high risk aversion also adjust more frequently than those with low risk aversion. c). It should not be a surprise to find that households with higher risk averse have less portion of their wealth in risky asset, as shown in column 7. Also it should not be a surprise to see those with very low risk aversion would like to borrow at the risk free rate and invest them in the risky asset. As a result, there are a few negative numbers for $A$ equals 0.9, 1.1 and 1.3. Note that even though higher risk averse households put more wealth in risk free asset, they do not necessarily also hold more in durable goods. For example, when $\lambda$ is equal to 0.01 and $A$ increases from 2.9 to 3.1, risky asset holding drops from 40.97 percent to 38.35 percent, and
at the same time risk free asset increases from 40.91 percent to 43.76 percent, but durable goods holding drops from 18.12 percent to 17.89 percent. From the next two categories, it becomes even more clear that durable share usually goes up for awhile and then goes down. This is because income effect and substitution effect have switched importance during the process. When risk aversion is low, substitution effect dominates income effect, so when money is taken out of risky asset, some of them are used to increase durable stock. However, when risk aversion is already high, income effect dominates substitution effect, and both risky asset and durable stock are reduced to increase risk free asset. It cannot be optimal for a household to keep increasing its durable stock share due to the high depreciation rate assumed in the above simulations. d). Last two columns show drift and standard deviation of income. It is obvious that more risk averse households hold less in risky asset, so they have lower drift and variation in income. Note that income is not exogenous in this model. The only source of income is asset returns. As a result, larger inaction bandwidth and higher drift and variation of income are direct result from lower risk aversion, it is not that households optimally choose higher bandwidth because of higher exogenous income drift and variation.

Next, I focus on the implication of higher adjustment cost. e). Higher adjustment cost entails larger bandwidth. Actually an analytical proof of this relationship can be found in Grossman and Laroque (1990) appendix. But what is truly important to households when facing higher adjustment cost is that they will try to reduce...
the frequency of adjusting. This is strongly supported by Column 4 in the table: expected time lengths between purchases are uniformly longer when adjustment cost is higher. Bandwidths are larger too, but it is secondary because it is nothing but a measure taken to reduce adjustment frequency. If adjustment frequency is reduced due to higher adjustment cost, household has to hold the durable stock for longer period of time. In order to reduce the average deviation between the actual durable stock and the optimal durable stock, they must now hold a larger durable stock. This is also supported by the result reported in Column 5: \( \left( \frac{K}{W} \right)^* \) is uniformly larger when \( \lambda \) is larger. g. Carefully examining the last 4 columns, I find that when adjustment cost is higher, the ratio between shares of risk free asset and risky asset is getting (slightly) bigger. That is equal to say, when \( \lambda \) is larger, besides that households put more wealth into durable stock, they also put more into risk free asset than in risky asset. Since CRRA utility function is used in the model, this should not be due to the wealth effect (i.e., higher \( \lambda \) makes people poorer in general, and then more risk averse.) It is not clear why this happens, but to look from the opposite perspective, these differences are almost negligible given such big changes in \( \lambda \). So I can also use this finding to roughly conclude that the proportions of risk free asset to risky asset are maintained by the household regardless of adjustment costs.
2.3 Ergodic Distribution

For each individual household across time, the state variable $y$ will evolve as equation (2.14) with boundary limits $y_1$ and $y_2$. Geometric Brownian motion can be viewed approximately as random walk with drift.\(^6\) That is equal to say, within the boundaries, infinitesimally small time elapse causes $y$ to change as a binomial distribution. Due to the Markov property of this process, the ergodic, or long-run, distribution of $y$ will be independent of the starting value of $y$. Intuitively, if we randomly pick the observation times which are separated by enough long time, the probability of observing each value of $y$ is a constant. In appendix B, I formally derive this ergodic distribution following Bertola and Caballero (1990) Dixit and Pindyck (1994), and Eberly (1994). This property holds at the micro level no matter there is aggregate shock or not, because each household only cares about the total amount of risk and does not distinguish between aggregate shock and idiosyncratic shock. However, this individual ergodic distribution is also the cross-section distribution of the durable stock if we assume all households are the same except for their current holdings of the durable stock and there are no aggregate shocks. This is a simple application of the Glivenko-Cantelli Theorem (see, e.g., Billingsley 1986). The intuition is, when a large number of identical households are observed and each one of them converges to identical stationary ergodic distributions, the

\(^6\) See Dixit and Pindyck (1994) Chapter 3 and 4, for example.
probability of observing one household having $y$ equal to $y'$ is the same as observing the same proportion of population having $y$ equal to $y'$. This is the same idea as the classical binomial distribution: if the probability of observing a Head is $p$ for any single coin toss, then this $p$ is also the proportion of Head observed either by tossing this coin a large number of times, or by tossing a large number of identical coins.

Figure 1 is an example of the ergodic distribution.\(^7\) Parameters are set at the same values as those used in Grossman and Laroque (1990) and Eberly (1994). Formal derivation of this distribution is in appendix B. This shape is very intuitive: it is piece-wise exponential with the highest density at the internal return point, and smoothly decreases to zero towards both boundaries. The internal return point should have the highest density because it can be reached not only from its own neighborhood but also from both boundaries. The density is also skewed toward the right because of the positive drift in $y$, which is caused by the positive return on financial wealth and the depreciation of durable stock. However, this shape is slightly different from what is reported in Eberly (1994). In figure 1, when $y$ is above the internal return point and below the upper bound, the density decreases slightly

---

\(^7\) Note this shape is slightly different from what I use in chapter 4. The reason is, the current ergodic distribution is derived for a general form of Brownian motion in which both the instantaneous drift and standard deviation are functions of the state variable, while in chapter 4 I only use the simple case where both the instantaneous drift and the standard deviation are constant. As a result, the former has an additional Jacobian function in it. However, since only the simple case has an analytical functional form, it is the natural choice for estimation.
in a convex shape and then turn concave, while Eberly (1994) shows it as always concave on the right hand side of the internal return point. Interestingly, this shape matches Eberly’s simulation result better than her own ergodic distribution. This difference is not crucial and is explained in detail in appendix B.
Chapter 3

MODEL WITH THRESHOLD

AUTO-REGRESSION ESTIMATION

In this chapter and the next chapter, I present two methods to empirically investigate the theoretical model of the previous chapter. I will first discuss the threshold auto-regression model, and then the mixture distribution model.

Household decision on durable goods replacement depends only on the current stock of the durable goods and nothing earlier. If the stock value goes beyond either threshold, household will execute adjustment and return the stock value to some preset target level. This idea suggests a threshold auto-regression model.
3.1 Model Specification

The threshold auto-regression model is specified as follows:

\[ y_{t+1} = y_t e^{\xi + \epsilon_{t+1}} \quad \text{if } y_t \text{ is within the bounds} \quad (3.1) \]

\[ y_{t+1} = \mu \quad \text{if } y_t \text{ is out of either bound} \quad (3.2) \]

Equation (3.1) holds when no adjustment occurs, the percentage change in \( y \) has a trend as well as a stochastic component. \( \xi \) is equal to the trend in lifetime wealth plus the depreciation rate, and \( \epsilon \) summarizes the total risk faced by the household, it includes both aggregate risk and idiosyncratic risk. \( \epsilon \) has mean zero and variance \( \sigma_\epsilon^2 \). \( \mu \) is the internal return point, same as \( y^* \) in the previous chapter. When \( y \) is out of either bound the household will adjust back to the internal return point, as specified in equation (3.2).

Take natural logarithm on both sides of equations (3.1) and (3.2), I obtain:

\[ \ln(y_{t+1}) = \xi + \ln y_t + \epsilon_{t+1} \quad \text{if } y_t \text{ is within the bounds} \quad (3.3) \]

\[ \ln(y_{t+1}) = \ln(\mu) \quad \text{if } y_t \text{ is outside of either bound} \quad (3.4) \]

If there is no human error, equations (3.3) and (3.4) will be followed exactly. However, there is other randomness in reality. There are at least two common mistakes that prompt me to add an error term to the right hand side of both equations. First of all, there is usually a difference between the time at which the household makes the decision and the time at which the survey unit records the decision. For
example, the survey unit probably records the purchase half a year after the household actually bought the automobile. So the recorded $y$ is not actually the internal return point, but plus some changes incurred by time elapse. Second, when the household reports its income, net worth and moving vehicle stock, it may overstate or understate the values for various reasons. Thus, let's summarize these into one error term $\nu$, it has mean zero and variance $\sigma^2$.

Now the model used in the estimation is:

\[
\begin{align*}
\ln(y_{t+1}) &= \xi + \ln(y_t) + \epsilon_{t+1} + \nu_{t+1} \quad \text{if } y_t \text{ is within the bounds} \tag{3.5} \\
\ln(y_{t+1}) &= \ln(\bar{y}) + \nu_{t+1} \quad \text{if } y_t \text{ is out of either bound} \tag{3.6}
\end{align*}
\]

Equations (3.5) and (3.6) form the basic estimation. To simplify the estimation, I also assume that $\epsilon$ and $\nu$ are not correlated with each other. As a result, the size of total uncertainty faced by the household can be distinguished from human error because the former only shows up in the first equation. Roughly speaking, the difference between errors from the first equation and that from the second equation gives me an idea on how large the total uncertainty is.

Let $(u, l)$ be the upper and lower thresholds to be estimated. Define:

\[
SSR(u, l) = \min_{(\xi, \bar{y})} \sum_{i=1}^{T-1} \{1 \{u > \ln(y_i) \} \{\ln(y_{t+1}) - \xi - \ln(y_i)\} + 1 \{\ln(y_i) > u \text{ or } \ln(y_i) < l\} \{\ln(y_{t+1}) - \ln(\bar{y})\}^2 \tag{3.7}\]

where $1\{.\}$ is the indicator function, which takes value 1 when the conditions in the parenthesis is satisfied and 0 otherwise. $T$ is the total number of observations per
individual. Then,

\[ (\hat{u}, \hat{l}) = \text{argmin}(SSR(u, l)) \]  

(3.8)

where a circumflex indicates the estimate for the parameter. Finally, estimates from equation (3.8) are treated as the true parameters and plugged back into equation (3.7), then \((\xi, \ln(\tau))\) as well as \((\sigma_x^2, \sigma_y^2)\) can be estimated using OLS. This algorithm is used on each individual household to recover its underlying parameters.

3.2 Data

The panel data set used here is from National Longitudinal Survey of Youth 1979 Cohort (NLSY79). It is an expansive survey sponsored by the U.S. Bureau of Labor Statistics, U.S. Department of Labor and conducted by the National Opinion Research Center (NORC) at the University of Chicago for the Center for Human Resource Research (CHRR) at the Ohio State University. The primary purpose of the survey has been the collection of data on labor force experience, labor market attachment, and investment in education and training by each respondent. However, the actual content of the survey is much broader than the above list. It also includes detailed records on respondent's income, assets (including moving vehicles) and demographic characteristics. Specifically, it provides a cross-sectional sub-sample of 6111 youths designed to be representative of the non-institutionalized civilian segment of young people living in the U.S. in 1979 and born between January 1,
1957 and December 31, 1964. The panel of durable stock holdings is constructed from this cross-sectional sub-sample, and the time period is 1987-1994.$^1$

Basic series used include total wealth, durable stock and a vector of demographic variables (denoted $Z$ in the model). Total wealth is net worth plus lifetime labor income. Following Eberly (1994), lifetime income is the present value of current and future income until age 66, discounted at an annual rate of 5 percent. All income and wealth are real.$^2$ Also following Eberly (1994), the durable stock is constructed as the moving vehicle stock times 10, since on average moving vehicle stock is 10 percent of the total durable stock (including housing). All moving vehicle stocks are real too.$^3$ Demographic variables included are:

- Age of the head of the household;
- Race of the head of the household;
- Sex of the head of the household;
- Marriage status of the head of the household;
- Education attainment of the head of the household;

$^1$Among the 6111 individuals, only about a quarter of them provide enough information for the estimation. See appendix C for detailed processing information of the data, including checking for sample selection problems.

$^2$Deflated by CPI-U. 1982-1984=100.

$^3$Deflated by CPI Transportation. 1982-1984=100.
• Number of adults living in the household;

• Number of children living in the household;

• License fee charged in the residence state of the household.

Sample statistics is reported in table 2. See appendix C for detailed information on the data processing.

### 3.3 Estimation Results

Table 3 is the summary of the result from the above estimation. As the findings of Lam (1991) and Eberly (1994), there exists significant heterogeneity across individuals. This is evident in variation of upper bound, internal return point and lower bound. In addition, the cross-sectional distribution of boundaries and internal return points is not symmetric, most of them are skewed to the right.\(^4\) To account for this feature of the estimates, I also report quartiles in table 3. Despite substantial heterogeneity, the estimates of the thresholds are in general compatible with the prediction of the theoretical model and the empirical results from other papers. For example, with the parameter values assumed in chapter 2 and risk aversion \(A = 1.1\) and adjustment cost ratio \(\lambda = 0.05\), the theoretical model predicts the lower bound, \(^4\) Lam (1991) uses an exponential scale factor on the boundary conditions, this also implies significant skewness in the cross section distribution of the actual boundaries.
internal return point and upper bound to be (1.19, 2.02, 2.84).\(^5\) An internal return point of 2.02 is very close to the median estimate (1.97). An upper bound of 2.84 is larger than the median estimate (2.68) but easily lies below the third quartile of the estimate (3.21). A lower bound of 1.19 is less than median estimate, and is also below the first quartile of the estimate (1.27). This is probably due to the fact that downward adjustment is rather unusual, and I use the observed lowest state variable value to be the proxy of it. This certainly overestimates the lower bound. \(\xi\) is usually believed to be around 0.12, which is 10 percent annual depreciation of durable goods and 2 percent annual aggregate drift. The mean of the estimates is 0.19, and the variation across household is substantial. This shows the drift of the state variable differs a lot across household, and it implies another side of the heterogeneity. The last two rows are estimates on the variation terms.

### 3.4 Determination of Bandwidth

Bandwidths are determined by individual characteristics. In the theoretical model of chapter 2, it depends on the constant risk aversion coefficient \(\lambda\), the adjustment cost coefficient \(\lambda\), the durable goods depreciation rate \(\alpha\), and other environment parameters. However, the theoretical model in chapter 2 doesn't model labor income as a source of exogenous uncertainty, it instead only considers the risk in a common

\(^5\) They are natural logarithm of (3.3, 7.57, 17.06).
stock market. This can be understood as the extreme case where labor income uncertainty is completely diversifiable. In general, uncertainty associated with labor income will also affect bandwidth. Specifically, it should be optimal for individual with higher labor income uncertainty to have larger bandwidth, because he should try to avoid adjusting right after experiencing good luck (bad luck) that could easily be reversed in the near future. To summarize, we expect that an individual with higher income uncertainty, higher drift in wealth and facing higher adjustment cost to have a larger bandwidth.

In order to check the validity of these implications, I regress the estimated bandwidths on average lifetime wealth, average drift of lifetime income, standard deviation of growth rate of lifetime income, and a proxy for adjustment costs. This proxy is the state license fee from The Statistical Abstract of United States, various years. It is calculated by dividing total license fee revenue by total number of registrations for each state and each year and averaged over the eight years. The auxiliary survey data from NLSY79 Geocode is used to identify each respondent's state of residence for each year. Then the state of residence is taken to be the one at which the respondent stayed the longest. Most respondents didn't move between states during 1987-1994. To summarize, the regression is as follows:

\[
\text{bandwidth} = \text{const.} + \varphi_1 \ln(\text{alw}) + \varphi_2(\text{drift}) + \varphi_3(\text{std.}) + \varphi_4(\text{lf}) \quad (3.9)
\]

Definition and calculation of the above variables are as follows:
• bandwidth is calculated as $(y_2 - y_1)/y^*$. 

• alw stands for average lifetime wealth, it is the simple average of lifetime wealth over 1987-1994.

• drift is the average growth rate of lifetime income over 1987-1994.

• std. is the standard deviation of lifetime income growth rate over 1987-1994.

• lf is the state license registration fee.

The regression result of equation (3.9) is reported in table 4.

The results are in general supportive of the theoretical model. First of all, $\varphi_1$ is not significantly different from zero, indicating that lifetime wealth level doesn’t have statistically significant effect on bandwidth choice. This can be used to support the assumption that liquidity constraint is not very important in households’ durable consumption decision. Eberly (1994) shows that income or wealth level should only have significant (negative) effect on bandwidth choice for the liquidity-constrained households but not for the non-constrained households. This is because when the liquidity constraint is binding, households can not transfer future income to finance current purchases, so upward adjustment ability is limited. This causes larger bandwidth than they would have optimally chosen. With higher income or wealth level, this constraint is less stringent and it should be reflected by a smaller bandwidth. Income drift and variation both have positive effects on bandwidth,
well as expected. However, the effect is only significant for variance (p-value=0.017) but not for drift. Eberly (1994) also report that the effect of lifetime income variation seems to be much more significant than its drift. This is probably due to measurement error in lifetime income. As is shown in the data appendix C, lifetime income is constructed from an estimate of predictable future income. If however the households use other “rule of thumb” to approximate the income perspective, our estimates may be not so accurate. With higher variation of lifetime income, households know they should not be fooled into a hasty adjustment and they need more time to check the permanence of the change. This warrants larger bandwidth. A notable success over previous evidence is the estimate of $\varphi_4$, the effect of license fee on bandwidth (p-value=0.009). License fee is a fixed adjustment cost associated with adjusting one's moving vehicle stock.\footnote{There is a subtle difference between adjustment cost here and the one implied by the theoretical model. In the theoretical model, the adjustment cost is in the form of irreversibility, i.e., it only takes the form of capital loss when resell a durable good. However, adjustment costs can take other forms too. It is very hard to find a proxy for $\lambda$ used in the theoretical model.} During 1987-1994, license fees across United States have a mean of $57/per registration and standard deviation of $18. Our results strongly support that respondents who reside in a state with higher license fee have larger bandwidths.\footnote{Eberly (1994) estimate is an insignificant negative number for the non-constrained group.}

A final note about this result is that I've maintained the assumption that people adopt constant threshold rules over this period of time. That is equal to say, $y_1$, 

$y_2$ and $y^*$ are all constants for every respondent. This is certainly not the case in general. Respondents may change their threshold rules due to significant changes to their demographic features, for example, an unexpected promotion, getting married or having a newborn baby. Moreover, households' altitude toward risk may change too. All of these may cause them to change their threshold rules. In order to capture this side of the story, we need much longer time series (at least 10-20 years of data for each family). This problem is partially resolved by exploiting the long-run stationary distribution of durable stocks. That is the main topic in chapter 4.

### 3.5 Empirical Cross-Section Distribution of Durable Stocks

In this section, I close the empirical evidence with a summary on the cross section distribution of durable stock holdings. In particular, I first compare the empirical cross sections distributions of $y$ across time. And then the theoretical ergodic distribution is compared with the empirical cross section to see whether they bear any resemblance.

Figure 2 is a 3-D picture of the empirical cross section distribution of $y$ over 1987-1994.\(^8\) It is obvious that the cross section distribution does not change drastically

---

\(^8\) Empirical discrete cross section distribution is first smoothed using Gaussian Kernel with
during this time period. However, it does change modestly in a way that is in accordance with the theoretical prediction of chapter 2. For example, from 1987 to 1988 and then to 1989, the distribution becomes spiky each year. This is because positive aggregate shocks synchronize individual behavior: more people choose to adjust their durable stock due to advancement of purchase under positive aggregate shocks, so more people are located at the center of the distribution. However, when the 1990-1991 recession comes in, idiosyncratic shocks replace aggregate shocks to exert the dominant effect. People have no incentive to advance adjustment decision; on the contrary, they may postpone their purchase decision under negative aggregate shocks. So idiosyncratic shocks spread people across the distribution and make the cross section flatter. This is why in 1990 the distribution becomes less spiky than 1989, and it stays like that until 1992. After 1992, it becomes spiky again when the economy starts to recover. This cyclical behavior of cross section distribution is in accordance with the theoretical prediction in chapter 2.

Another feature is that the cross section distribution has a much longer tail than the ergodic distribution. Even though these empirical cross section distributions are similar to the ones shown in Eberly (1994), they are unlikely to come from a single ergodic distribution. Unlike Eberly's result, a simple Komogrov-Smirov goodness of fit test (two-sided) clearly rejects the null hypothesis of same distribution for all eight optimal bandwidth from B. W. Silverman (1986). Top 5% of the observations are dropped to make the picture more compact. Also note 1991 values are constructed.
empirical cross sections. Given the sample size I use, the critical value of \((p > 80\%)\) should be around 0.028. Recall that the lower bound of the ergodic distribution is 3.30, and the upper bound is 16.9. Table 5 shows the minimal requirement on the ergodic thresholds.

Table 5 clearly shows that in order to pass the Komogorov-Smirnov goodness of fit test at \((p > 80\%)\) critical value, the ergodic distribution can not have a lower bound above 1.45-1.97 range, or an upper bound below 63.97-86.94. Check table 1 in chapter 2, none of the simulated lower bounds and upper bounds satisfies the above requirements. To put in another way, in order to claim that each household has the same ergodic distribution that also resembles the ergodic distribution like Eberly (1994) does, I need the ergodic distribution has a lower bound at around 1.45-1.97 and at the same time a upper bound at around 63.97-86.94. Apparently none of the simulated ergodic distribution comes even close to this range. As a result, we can easily reject the hypothesis that the cross section distribution is the same as the ergodic distribution for all 8 years.

---

9 I seriously doubt Eberly's test result, because by looking at the figures in her paper, they should not have passed the Komogorov-Smirnov test. For example, given the sample size she uses, the critical value for \(p > 80\%\) is around 0.04. But the tail of the empirical distribution beyond the ergodic distribution is at least 0.15. So based on this "eyeball" check, the empirical distribution should not have passed the test.

10 That means if the null hypothesis is true, the probability of false rejection is less than 20%. Higher \(p\) leads to more stringent requirement for threshold levels, so stronger rejections. Note that the Komogorov-Smirnov test statistic should be the maximum distance between the CDFs of two distributions. Here for simplicity I only use the CDFs at the boundary of the ergodic distribution thresholds. Clearly they only provide sufficient condition for rejections, not necessary conditions. However, a sufficient condition is enough to illustrate the point at this moment.
This rejection should not be much of a surprise because there is substantial heterogeneity in the cross section. It is more reasonable to think that different households may have different ergodic distributions and the observed cross section consists of individual distribution with different boundaries. As a matter of fact, this idea is confirmed by simple sample statistics. Using the data on household state variable $y$ for each year, I calculated the within household variation of $y$ and across household variation of $y$. If all household assumes identical ergodic distribution, the within household variation should be the same as the across household variation. Otherwise, it is more reasonable to believe that the cross section is the result of heterogeneous individual ergodic distribution. It turns out that the average standard deviation within household is around 13, while the standard deviation of the average value of the state variable across household is around 20. That is equal to say, the variation between households are much larger than that can be accounted for by within variation. This supports that it is plausible to believe heterogeneous households whose ergodic distributions have different support ranges from the cross section distribution.

Based on the above analysis, it would be too stringent and misleading to follow any one of the ergodic distributions listed in table 1. In the next chapter, I am going to adopt another model to take into account heterogeneous households.
Chapter 4

MODEL WITH MAXIMUM LIKELIHOOD ESTIMATION

Following Grossman and Laroque (1990), I model a household’s holding of durable goods as adjusting at discrete periods in time in response to a single state variable

\[ \bar{y}_t = \ln \left( \frac{W_t}{K_t} \right) \]

where \( W_t \) is household’s lifetime wealth at time \( t \), and \( K_t \) is its durable goods stock at time \( t \).\(^1\)

It is quite intuitive to think that a rational household would like to keep a roughly constant proportion of its lifetime wealth as durable goods. However, lifetime wealth

\(^1\)Note the definition of \( \bar{y} \). It is the logarithm of the wealth-durables ratio, not the share of durable goods in wealth. I adopt this definition from Grossman and Laroque (1990) and Eberly (1994) to insure comparability. In addition, the definition of \( \bar{y} \) is slightly different in Grossman and Laroque (1990) and Eberly (1994). The former defines \( y \) as \( \frac{W}{K} - \lambda \), while \( \lambda \) is the size of adjustment cost measured as a percentage of durable stock. I follow the latter and assume \( \lambda \) is zero. Also, I take the logarithm of it to simplify the model exposition.
usually subjects to random shocks such as job changes or stock market fluctuations. In addition, the value of durable goods stock may depreciate or appreciate, too. All of the above implies that the state variable $\bar{y}$ may have very complicated dynamics. Most importantly, because of non-convex adjustment costs, it is not optimal for the household to adjust durable goods stock whenever $\bar{y}$ deviates from the desired level. Instead, the household adopts an (S,s) rule which has an internal return point $\tilde{r}$, an upper trigger $ut$ and a lower trigger $lt$.\(^2\) The household will only adjust the durable stock to restore $\bar{y}$ to $\tilde{r}$ if either trigger is reached. Take the stock of automobiles as an example. Most households would experience an increasing $\bar{y}$ over time due to wealth growth and automobile depreciation. When this pushes $\bar{y}$ to the upper trigger $ut$, the household purchases automobiles in order to restore $\bar{y}$ to $\tilde{r}$.\(^3\) Occasionally some households may suffer a large wealth loss. In that case, if $\bar{y}$ drops to the lower trigger $lt$, the household will sell its automobile and buy a cheaper one, restoring the value of $\bar{y}$ to $\tilde{r}$, too.\(^4\)

---

\(^2\) The optimality of (S,s) rule is rigorously derived in Grossman and Laroque (1990) under very restrictive conditions.

\(^3\) In practice, most households would trade in an old automobile for a newer, higher-valued one.

\(^4\) Generally speaking, the return point from reaching the upper trigger may be different from that of reaching the lower trigger. In this paper I only consider the simple case of one return point. This is true if the adjustment cost is proportional to the current stock of durable goods.
Specifically, I assume the state variable $\bar{y}$ follows the simplest version of Itô processes:

$$d\bar{y} = \bar{\alpha} dt + \bar{\sigma} dz$$  \hspace{1cm} (4.1)

where $d\bar{y}$ is the change in $\bar{y}$ in the instant of time between $t$ and $t + dt$, $dz$ is the increment of standard Brownian motion, and $\bar{\alpha}$ and $\bar{\sigma}$ are the instantaneous drift and standard deviation, respectively.\(^5\) That is to day, the percentage change in $\frac{W}{K}$ has mean $\bar{\alpha}$ and standard deviation $\bar{\sigma}$.

Equation (4.1) is a diffusion process. Without boundary (return) conditions, it will diverge. I assume there exist three control values of $\bar{y}$ which defines the (S,s) rule: the internal return point (target) $\bar{r}$, the upper trigger $ut$, and the lower trigger $lt$. Whenever $\bar{y}$ is driven to either trigger point, it will be returned to the target and the process continues. To simplify the model, I further assume that both bands are of the same size and equal to $bw$. That is,

$$ut - \bar{r} = \bar{r} - lt = bw$$  \hspace{1cm} (4.2)

This turns out to be crucial for the estimation program to converge in a reasonable time. In order to justify this practice, I checked the numerically solved band widths in table 1. All 36 pairs of upper band and lower band are very close to each other.

\(^5\) The Itô process derived in Grossman and Laroque (1990) takes the general version of Itô processes. Here I use the simple version so that an analytical ergodic distribution function can be derived.
The difference between bands has a mean 0.05 and standard deviation 0.06, and upper band and lower band has a correlation coefficient as high as 0.98. All of these indicate that this is a reasonable simplification. Note that this assumption does not imply equal band width measured as share of durable stock in total wealth. It turns out that usually upper band is larger than lower band in such form.

Equation (4.1) and boundary conditions \((i\bar{r}, bw)\) constitute a controlled diffusion process. Figure 3 plots a sample path of this type of process. Consider \(\bar{\sigma} = 0.11\), \(\tilde{\sigma} = 0.59\), \(i\bar{r} = 2\), and \(bw = 1\), the state variable \(\tilde{y}\) starts from the internal return point and wanders within the band.\(^6\) Once in a while it will hit the upper band and then start again from the internal return point. Occasionally it hits the lower band. An adjustment is indicated by a circle in the figure. In this particular realization, the first four adjustments are from hitting the upper band, only the last one is from hitting the lower band. A total of five adjustments have occurred during this 20 years of time. Intuitively, this process can be thought of as a random walk with reflective barriers. The influence of initial value of \(\tilde{y}\) will disappear once either barrier is hit, because of the Markov property. As the motion goes back and forth trapped between the two barriers, we expect it to settle down to a stationary long-run (ergodic) process.\(^7\) That is to say, there is a stationary long run distribution for

\(^6\) The continuous process is first discretized into the corresponding random walk with a drift, according to Dixit and Pindyck (1994). Step length is a month.

\(^7\) See Dixit and Pindyck (1994) page 83-84 for more details.
the state variable $\bar{y}$, and the functional form of the distribution will depend on $\alpha$, $\sigma$, as well as boundary condition $(\tilde{\tau}, bw)$. Appendix B gives the formal derivation of the general form of this ergodic distribution function. Alternatively, the same result can be proved by applying Kolmogorov forward equation. Here I only rewrite the pdf of this distribution, simplified by the assumption that the upper band width and the lower band width are of the same size:

$$
\phi(\bar{y}) = \frac{(e^{\gamma \cdot bw} - 1)(e^{\gamma \cdot (\tilde{\tau} - \tilde{\tau})} - e^{-\gamma \cdot bw})}{bw \cdot (e^{\gamma \cdot bw} + e^{-\gamma \cdot bw} - 2)} \quad \text{if } \bar{y} \in [\tilde{\tau} - bw, \tilde{\tau}] \quad (4.3)
$$

$$
\phi(\bar{y}) = \frac{(1 - e^{-\gamma \cdot bw})(e^{\gamma \cdot bw} - e^{-\gamma \cdot (\tilde{\tau} - \tilde{\tau})})}{bw \cdot (e^{\gamma \cdot bw} + e^{-\gamma \cdot bw} - 2)} \quad \text{if } \bar{y} \in [\tilde{\tau}, \tilde{\tau} + bw] \quad (4.4)
$$

where $\gamma$ is defined as $\frac{\alpha}{\sigma}$.

Figure 4 a-f show some sample ergodic distributions based on $\tilde{\tau} = 2$, $bw = 1$ and various values of $\gamma$. The shape of the ergodic distribution is very intuitive. It is piece-wise exponential, increasing from zero at the lower trigger point to the peak at the target, and decreasing from there to zero again at the upper trigger point. This is because households turn to spend more time at the neighborhood of the target, and the probability decreases as they wander towards either trigger point. Positive $\gamma$ drives the distribution mass to the right, while negative $\gamma$ drives it to the left. When $\gamma$ equal to zero, it degenerates to a symmetric triangular distribution.

---

8Mainly because that durable stock is exponentially depreciating.
When γ approaches positive or negative infinity, the distribution becomes uniform on the upper band side or lower band side, respectively. Bertola and Caballero (1990) shows a series of graphs similar to figure 4 here.

Based on the definition of $\bar{y}$, it should always be positive because durable stock $K$ is part of total wealth $W$. There are several ways to imbed this notion into the model. Truncated distribution usually has complex analytical form and proves to be ominous for estimation. Instead, I deliberately choose the following function forms for the boundary condition determination:

$$\bar{t}_r = \exp(Z \cdot \beta^t_r + \epsilon^t_r) \quad (4.5)$$
$$bw = \exp(Z \cdot \beta^{bw} + \epsilon^{bw}) \quad (4.6)$$

where $Z$ is a vector of demographic variables, $\beta$ is the corresponding vector of parameters to be estimated. $\epsilon^t_r$ and $\epsilon^{bw}$ are the random disturbances to the target and band width, respectively. That is to say, the total heterogeneity in household choice of boundaries is decomposed into an observable part $Z \cdot \beta$ and an unobservable part $\epsilon$. I assume the $\epsilon$s follow joint normal distribution with correlation coefficient $\rho$.

The choice of exponential function form guarantees that the target point must be positive, and the upper trigger point must be positive too. Even though it does not guarantee the lower trigger point to be positive, it turns out in later estimation that it is unimportant: in the estimated model the probability of observing a negative lower trigger is virtually zero. Most importantly, I circumvent the issue of
distribution truncation and both random shocks in the model now have complete distributions. This will simplify both parameter interpretation in chapter 5 and numerical simulation in chapter 6.

The above model construction has a very intuitive interpretation. Each observation of $\tilde{y}$ can be viewed as being generated by a two step procedure. First, a pair of random shocks to $\tilde{r}$ and $bw$ is drawn. Combined with the household's demographic features, it determines the ergodic distribution of this household. Second, a random value is drawn from this distribution, and that is the observed $\tilde{y}$. That is to say, $\tilde{y}$ follows a mixture distribution. Since it is unclear which exact ergodic distribution each $\tilde{y}$ is drawn from, I need to integrate over all possible values of $\epsilon$, i.e., all unobserved heterogeneity needs to be integrated out. Intuitively, the likelihood for each observation should reflect the mean probability of observing that certain $\tilde{y}$ after considering all possible values of $\epsilon$. According to this logic, the likelihood function can be written as:

$$L(\tilde{y}|\beta, \gamma, \sigma_{bw}, \sigma_{\tilde{r}}, \rho) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\tilde{y}|e^{bw}, e^{\tilde{r}}) f(e^{bw}, e^{\tilde{r}}) de^{bw} de^{\tilde{r}}$$

(4.7)

where $f(., .)$ is the joint pdf of $e^{bw}$ and $e^{\tilde{r}}$.

I choose the parameters to maximize the joint (log) probability of observing the whole collection of $\tilde{y}s$. That is:

$$\{\hat{\beta}, \hat{\gamma}, \hat{\sigma}_{bw}, \hat{\sigma}_{\tilde{r}}, \hat{\rho}\} = \text{argmax}\{\sum_{i=1}^{n} \log(L(\tilde{y}_i|\beta, \gamma, \sigma_{bw}, \sigma_{\tilde{r}}, \rho))\}$$

(4.8)
Note that even for the same household, this ergodic distribution is usually not the same from one period to another. This is true not only because of observed changes in the demographic characteristics, but also because of unobserved random shocks to the boundary conditions. That is to say, es are iid in both cross-section and time series. In this sense, this model has a dynamic aspect which can be further exploited in constructing aggregate time series.

Estimation of the above model seems straightforward, but preliminary regression revealed some problems which deserves special notice. First of all, due to the strong left-ward skewness observed in the sample (See Figure 5-a), the estimated $\gamma$ is around -2. This certainly contradicts common knowledge because we would expect the state variable $\dd$ to drift upward due to growth in wealth and depreciation in durables. As a result, we have strong priori belief that $\gamma$ should be positive. In order to accommodate this priori belief, I set $\gamma$ at the sample analog values calculated as follows. First, I take the first difference of eight observations for each individual, this gives all changes in $\dd$. However, some of the changes are due to the adjustment and they do not reflect the diffusion process in equation (4.1). Obviously these "jumps" usually are much larger than regular changes. In addition, common knowledge implies that a typical household would adjust once or twice during this eight year

---

9 In order to construct $\dd$, I need to construct wealth $W$ and durable stock $K$. This is discussed in detail in appendix C. In short, wealth is the sum of net worth, including durable stock, plus human capital. Durable stock is taken as ten times automobile stock, which is directly available in the data.
span. As a result, I tried removing the largest one and largest two differences from each household, and treat the rest as a sample purely from the diffusion process. Without such sample cleaning, the sample of first differences has a median of 0.10 and standard deviation of 0.79. After dropping the largest one for each household, they become 0.11 and 0.55 respectively. Finally, after dropping the largest two, they become 0.11 and 0.39. I believe the latter two are more plausible, so I use both in the estimation. By the definition of $\gamma$, it is set at 0.74 and 1.37, respectively. It will become clear soon that they do not produce dramatically different estimates. In addition, both of them does not produce very different estimates from setting $\gamma$ equal to -2.$^{10}$

In addition, preliminary estimation also shows that the correlation coefficient $\rho$ strongly converges to 1, which implies that households with higher targets always have larger band widths. So in the following regression, $\rho$ is set equal to 1.$^{11}$

---

$^{10}$ This point is further explained when estimation results are presented in the next chapter. Positive $\gamma$ is also crucial for carrying out the macro dynamic simulation in chapter 6.

$^{11}$ This is probably caused by the fact that the sample histogram shows skewness to the right with a thick tail. The model described in equation (4.5) and (4.6) is log-normal and usually has thin tail. In order to fit the thick tail, the estimates need to be able to generate more observations to the right of the peak of the sample distribution. This creates high correlation between target and band width. After all, in the direct model below, I find significant positive correlation between the two shocks, this indicates that setting $\rho$ equal to 1 may not be too far away from the true correlation.
I used the routine GAUSS MAXLIK (BFGS algorithm and STEPBT step search) to estimate the rest of the parameters. A Monte Carlo method is used for the numerical integration in equation (4.7), which is the most time-consuming part of the estimation.\textsuperscript{12}

The estimation results are reported and discussed in the next chapter.

\textsuperscript{12}In the Monte Carlo numerical integration, a random sample of size 1000 is drawn in each iteration. It takes roughly 4 days on a PIII 500 for the program to converge. A long list of included demographic variables also contributed to the slow convergence. Convergence becomes prohibitively slow without the equality assumption on band widths.
Chapter 5

ESTIMATION RESULTS

Data description is the same as in section 3.2.

5.1 Homogeneous or Heterogeneous?

Eberly (1994) concludes that her sample is drawn from a single deterministic ergodic distribution. However, this is unlikely to be true for my data set. Figure 5 shows four histograms for the panel data on $\bar{y}$. If every household has the same deterministic ergodic distribution, its band width should be at least as large as the support of the whole sample histogram. However, we can obtain a rough idea of each individual household's ergodic distribution by approximating its target and band width as follows. From the eight years of observations for each household,

\footnote{The data set she used is from Surveys of Consumer Finances conducted in 1983 and 1986.}
we can use half of the difference between the maximum and the minimum observed value as a proxy for the band width. In addition, we can use the mean or median of the eight observations as a proxy for the target. Figure 5-a is the whole sample histogram, and it obviously has a band width of at least three and a target around two, if the hypothesis of homogeneity is true. Figure 5-b is the histogram of approximated band width, which has mean 0.89 and standard deviation 0.43. Figures 5-c and 5-d are the histograms of the means and medians, respectively. In both cases the target has mean around 2 and standard deviation around 0.6. It is obvious from these figures that most of the households have band widths less than a third of the band width suggested by the whole sample histogram. Furthermore, both the target point and the band width vary a great deal across households. This evidence certainly militates against the hypothesis of homogeneity. A model that allows for heterogeneity may be able to explain the phenomena evident in Figures 5 a-d. The observed variation may come from observable differences in household characteristics and other unobservable differences. Finding out how each type of heterogeneity affects (S,s) rule is one of the main purposes of this work.

I estimate and compare the following three models:

- Direct Model: Treat the above proxy as the true target and band width for each household and estimate the parameters directly.

- Homogeneous model: Assume that every household has the same deterministic
ergodic distribution and estimate the parameters of the distribution.

- Heterogeneous Model: Assume that the household's ergodic distribution is affected by both observed and unobserved heterogeneity and estimate the mixture distribution model.

This comparison formalizes the critique of the homogeneity hypothesis offered above.

5.2 The Direct Model

An interesting question to ask at this early stage is, what if we simply take the proxies above as the true values? I estimate such a model in this subsection. Specifically, the system of equations to be estimated includes (4.5) and (4.6) only. Maximum likelihood is used to estimate the parameters. Table 6 reports the results with observed heterogeneity, while table 7 reports those without.

Taking the mean or the median as the target doesn't seem to matter much. In both cases demographic variables have shown small but significant effects on target and band width. Likelihood ratio tests clearly reject the null of no influence from demographic variables, but the inclusion of the vector of demographic variables only reduce the variation in $\sigma$s by a very small amount.

Other interesting findings include:

- Married people have smaller target and band width. This implies, ceteris
paribus, they hold more wealth in durable goods and adjust more often.

- Households with more children turn to have larger target and wider band, which implies that they hold less wealth in durable goods and adjust less often.

- Households reside in a state with high license fee turn to have larger target and wider band, which implies that they hold less wealth in durable goods and adjust less often.

- Cross section variation in target is positively correlated with that in band width.

Evaluated at the sample mean, the direct model (with mean as target) implies that target has a mean of 2.14 and standard deviation of 0.40, while band width has a mean of 0.89 and standard deviation of 0.18.² These results clearly indicate significant amount of heterogeneity in households' choices of target and band width.

Even though the direct model has produced some interesting insight of households' (S,s) choices, it is not an efficient model. Since it only uses four simple statistics from the sample, i.e., mean, median, maximum and minimum, it ignores information borne in each individual observation of \( \tilde{y} \). In the next two subsections

²These statistics are calculated by utilizing that if \( x \sim N(\mu, \sigma^2) \), then \( E(e^x) = e^{\mu + \frac{1}{2}\sigma^2} \) and \( V(e^x) = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2} \).
I shall turn to the new method introduced in chapter 4. I will first discuss the homogeneous case, and then the heterogeneous case. In both cases I will use the result from the direct model as reference.

5.3 The Homogeneous Model

The homogeneous model is a restricted version of the model in equation (4.5) and (4.6). We can set all other parameters, including the variance terms, equal to zero and only estimate the constant term. Because this model does not involve a mixture distribution, its estimation is straightforward. To maintain the comparability between models, I estimate the following form:

\[ \tilde{ir} = \exp(C_{ir}) \]  
\[ bw = \exp(C_{bw}) \]

And the likelihood function is:

\[ L(\tilde{y}|C_{ir}, C_{bw}) = \phi(\tilde{y}|C_{ir}, C_{bw}) \]

The set of parameters \((C_{ir}, C_{bw})\) is estimated by Maximum Likelihood as:

\[ \{C_{ir}, C_{bw}\} = \text{argmax}\{\sum_{i=1}^{n} \log(L(\tilde{y}_i|C_{ir}, C_{bw}))\} \]

Obviously, the above estimation is searching for the pair of \(C_{ir}\) and \(C_{bw}\) such that the probability of observing the whole sample is maximized. The estimation
result is reported in Table 8. For both cases of \( \gamma = 0.74 \) and \( \gamma = 1.37 \), the targets and band widths are precisely estimated. These values imply that the population target and band width is equal to 1.92 and 2.35 for the former case, and 1.57 and 2.70 for the latter one. These results are inconsistent with those from the direct estimation. In particular, the band width is about three times as large as the ones observed for most individuals.

In order to evaluate the goodness of fit with the sample, the estimated ergodic distribution is plotted with a kernel-smoothed sample histogram in Figure 6-a, and a cdf plot with Kolmogorov-Smirov test statistics is shown in Figure 6-d. Clearly both pdf and cdf plots show significant difference in these two distributions, the single homogeneous ergodic distribution does not match the sample histogram very well. In fact, the Kolmogorov-Smirov test statistic is 0.136, much larger than the 5 percent critical value of 0.013. As a result, I can easily reject that the sample is drawn from the estimated homogeneous ergodic distribution.

So far I have provided a lot of evidence against homogeneity in households’ (S,s) choices. All of them certainly suggest a heterogeneous model instead. In the next subsection, I shall focus on two questions:

- Does the heterogeneous model achieves a better fit of the sample?

---

3 For kernel smoothing I use a Gaussian kernel and optimal bin widths from Silverman (1990). Solid lines are for the observed sample, while dashed lines are for the simulated model. The 5 percent critical value of the Kolmogorov-Smirov test for a sample of 9429 is approximately 1.26 percent. See Daniel (1978) or Wand and Jones (1995) for detail.
• Do the demographic characteristics affect household's choice of target and band width in a way consistent with the existing theory?

5.4 Heterogeneous Model

In this subsection I estimate the full version model described in chapter 4. That is, the mixture distribution model with both observed and unobserved heterogeneity. This basically involves equations (4.3) to (4.8). Estimation result is reported in table 9. Most of the demographic variables affect both the target and band width in a small but statistically significant way. Unlike in the direct model, most of the demographic parameters are precisely estimated. In general, target point indicates (the inverse of) the desired level of durable goods stock as versus total wealth, so a higher target point implies that a smaller portion of wealth is durable goods. The band width indicates the range of inaction and is thus roughly an inverse measure of the frequency of adjustment. Households with larger band width will adjust durable stock less frequent than those with smaller band width. The results can be summarized as follows.

First of all, setting $\gamma$ to 1.37 does not produce dramatically different set of estimates from that of $\gamma$ equal to 0.74, so I will focus on the set of estimates for $\gamma$ equal to 0.74. Age doesn't seem to affect either the target or the band width in a statistically significant way. However the signs of the estimates do indicate that older
people may hold more moving vehicles and adjust more often. Attanasio (1999) reports that older people have a statistically smaller target durable-nondurable goods ratio, and band width is smaller too. Black households turn out to hold more moving vehicles and adjust more often, while Hispanic households hold less and adjust less often. Households with a male head hold more moving vehicles and adjust more often, a result similar to one of Attanasio's. Married people are also found to hold more moving vehicles and adjust more often. This is similar to the results from direct estimation. People with a college degree turn out to hold fewer moving vehicles and adjust more often. An increase in the number of adults in the household does seem to reduce the moving vehicles holdings (not statistically significant) and reduce the frequency of adjustment, another finding similar to one of Attanasio's. Households with more children have fewer moving vehicles and adjust less often, which is similar to that from the direct model, and also one of Attanasio's findings. Finally and most importantly, as a proxy for adjustment cost, households reside in a higher license fee state turn out to hold fewer moving vehicles and adjust less frequent. This finding is consistent with the theoretical prediction of Grossman and Laroque (1990). Both the direct model above and Attanasio (1999) have found the same result.4

4 All the above results are very similar to those estimated by setting $\gamma$ equal to -2 except the parameter for education in target determination. The result for $\gamma$ equal to -2 is not presented here but is available upon request.
Even though most of the demographic characteristics affect the target and band width highly significantly, their total explanatory power is rather small. Compared with a model with no demographic variables, as shown in table 10, the standard deviation of the unobservable heterogeneity in the target is reduced from 0.216 to 0.206, and that for band width is only reduced from 0.321 to 0.302. However, the LR tests strongly indicate that the demographic variables should be included.

In summary, the model performs very well in estimating the parameters of both observed and unobserved heterogeneity. In particular, the estimates on the proxy for adjustment costs, license fees, do indicate that households respond rationally to their environment. Moreover, the heterogeneous model has also achieved a much better fit than the homogeneous model. Figures 6-b and c plot the density comparison of the observed sample (kernel smoothed) with the simulated density from the estimated model. The densities clearly show strong resemblance to each other. Figures 6-e and f are the cdf plots with the Kolmogorov-Smirnov test statistic. Even though it still fails the test at a 5 percent critical value (by a small margin), the heterogeneous model obviously fits much better than the homogeneous model.\(^5\)

\(^5\) In order to improve the accuracy of the simulation, variations in demographic variables are also taken into account. First, sample correlation is calculated from all observations. Second, this correlation matrix, including that of the unobserved random shocks ($\sigma^{tr}$ and $\sigma^{br}$), is used in a multivariate normal random number generator to obtain 1000 pseudo households, then the population pdf or cdf is taken as the simple average of each individual component since I assume each draw is independent. Obviously the result will be affected by the seed of the random number generator which is automatically assigned by computer. In order to improve accuracy, I run 100 iterations and take the average of the distribution function and the test statistics. At such a large number of random draws in each iteration, the variation in the final results (such as test statistics) can be controlled at as low as 20 percent of the mean.
Comparing the estimated (mean) target and band width with those shown in figure 5, I find that both the target point and the band width match very well with the sample. The estimated target point has a mean of 2.00 (or 1.93 if $\gamma = 1.37$) and standard deviation 0.19 (same if $\gamma = 1.37$). The mean is very close to the statistics shown in Figure 5-c and d, and the variation is smaller. This is because the estimated target is calculated at the sample mean and all observed heterogeneity are ignored. As for band width, the estimated mean is 1.35 (or 1.29 if $\gamma = 1.37$), and standard deviation is 0.19 (same for $\gamma = 1$). The estimated band width is larger than the "observed" band width because in the set up of the model I assume that the difference between the maximum and minimum of the observed state variable only provides a lower bound for that measure. This result clearly re-emphasizes a crucial point mentioned in other studies that with limited (time series) observation, it is very difficult to estimate the band width, and underestimation is a common problem. However, the model here seems to be able to get around this problem.

To present the comparison of results in a more transparent way, I convert the estimates into an easy format. Table 11 lists the predicted (mean) boundary conditions measured in terms of percentage of wealth held in durable goods, i.e., $\frac{K}{W}$ instead of the original $\frac{W}{K}$. For example, when $\gamma = 0.74$, the mean target point is 2 and mean band width is 1.35. This implies the lower trigger and upper trigger is 0.65 and 3.35 respectively. Taking the inverse of the exponential of each one of them gives the boundary conditions measured as durable share in wealth. My model
predicts that households on average hold about 13.53 percent (or 14.51 percent for $\gamma = 1.37$) of their wealth in durable goods. This is very close to the value of 12.99 percent from both Eberly (1994) and Grossman and Laroque (1990). However, the inaction range is larger than both previous work: the lower trigger point is 3.51 percent (or 4.00 percent if $\gamma = 1.37$), while the other two are around 6 percent; the upper trigger point is more than 50 percent, while the other two are much smaller. However, Eberly (1994)'s upper trigger is very imprecise because only very few downward adjustment is observed, as a result it may grossly underestimate the size of the band width.

---

6In Grossman and Laroque (1990), this number is the numerical solution for a special set of parameter values (see page 45 Table V of that paper). Eberly (1994) takes the value of $y$ right after adjustment as target, so her approach is similar to the “observed band” model in Attanasio (1999).

7See Eberly (1994) page 419 Table 4.
Chapter 6

Macro Dynamics

The early literature using aggregate data found aggregate expenditure on durable goods to be rather smooth and persistent. For example, Mankiw (1982) showed that within the representative agent framework, durable expenditure should follow an ARMA(1,1) process if there is no adjustment cost. However, the moving average coefficient turned out not to be statistically significant, and he could not reject that the true underlying process is an AR(1) process, which is the same as Hall (1978) derived for its non-durable counterpart. In order to salvage the model, Bernanke (1984) introduced convex adjustment costs to explain the smooth and persistent adjustment. With convex adjustment costs, it can be shown that it is optimal for the representative agent to adjust her durable stock continuously and by small amounts. That is to say, the representative agent will only close a fraction of the gap between the desired and actual durable stock each period. However, this contradicts
the observation that most households adjust their durable stocks infrequently. For example, Lam (1991) reported that in the period of 1967-1970, 23 percent of households had zero expenditure in all three years, 27 percent had zero expenditure in two years, and only 12 percent had expenditures in all three years.\(^1\) This microeconomic observation is consistent with an \((S,s)\) model. In the presence of non-convex adjustment costs, households only adjust their durable stocks sporadically and by lumpy amounts. As a result, each of their durable goods expenditures should be positive only once in several years and zero otherwise. Unless the household expenditures are somehow fully synchronized, the aggregate behavior will differ from the individual behavior. As a result, the representative agent framework is not suitable for describing the macro dynamics of expenditure on consumer durables.

Another feature of aggregate durable expenditure is its high elasticity (i.e., over sensitivity) with respect to even transitory income shocks, and its slow movement back to the steady state level after these shocks. Figure 7 plots the quarterly real per capita expenditure on motor vehicles and parts as well as real per capita disposable income. Recessions are marked with vertical bars.\(^2\) It is obvious that the expenditure series is more volatile than the income series, and its recoveries from recessions are usually slow. For example, in the recession of 1981q3 to 1982q3,

\(^1\) Hendrick and Youman 1976 four-year panel study of consumer behavior.

\(^2\) Source: Bureau of Economic Analysis. 1959q1-1999q1. Chained 1996 dollar. RDPIPC stands for Real Disposable Personal Income Per Capita. REPEPC(M) stands for Real Durable Expenditure Per Capita on Motors and Parts.
real income dropped 2.6 percent and real durable expenditure dropped 8.2 percent, which implies an elasticity of at least 3. In the following recovery period of 1983q1-1984q1, income increased by 2.9 percent and real durable expenditure increased by 16 percent, which implies an elasticity greater than 5. This evidence casts doubt on the Permanent-Income/Life-Cycle Hypothesis (PIH/LCH) within the representative-agent framework.

In this chapter I use Monte Carlo simulation to check whether the model estimated in the previous chapter is capable of generating aggregate dynamics that match the above key features of the observed data. In particular, I check both the initial response and subsequent dynamics of aggregate expenditure after an aggregate shock to wealth. If the answer is yes, then we know that over-sensitivity observed in aggregate data is simply an aggregation phenomenon of heterogeneous (S,s) households, and certainly can not be used to reject PIH/LCH at the individual level. For comparison, I also present dynamics generated from three other models: a representative-agent model without adjustment costs (Mankiw), a representative-agent model with convex adjustment costs (Bernanke), and an (S,s) model with homogeneous households (Eberly).

Main features and findings of the simulation are summarized as follows:

- Each household follows an (S,s) rule, which is consistent with a modified version of PIH/LCH.
• The initial response of aggregate expenditure to the shock is highly elastic. By construction, this must be due to changes in the number of adjustments rather than the expenditure per adjustment. This finding resolves the puzzle of over-sensitivity.

• Shocks have long-lasting effects, due to the \((S,s)\) structure of individual adjustment. That is, some households respond to current shocks only in some later period when they reach the bounds.

• The simulated elasticity of initial response is much closer to the observed value than the reference models.

6.1 Simulation Setup

For simplicity, I used the estimation results from empirical model with only unobserved heterogeneity to generate pseudo households.\(^3\) I drew 10,000 households that differ from each other in target values and band widths. Without loss of generality, I normalize wealth to one. uncertainties will be for the state variable \(\bar{y}\), which is the logarithm of the ratio of wealth to the durable stock. In conformity with the empirical results reported in the previous chapter, I set the parameters \(\bar{\alpha}\) equal to 0.11, and \(\bar{\sigma}\) equal to 0.55 (in which case \(\gamma = 0.74\)) and 0.39 (in which case \(\gamma = 1.37\))

\(^3\) Simulation with both observed and unobserved heterogeneity yield similar results.
respectively. As a result, \( \frac{W}{K} \) trends upward 11 percent per year, and has standard deviation of 55 percent or 39 percent per year.\(^4\)

At the very beginning each household started at its target value. The drift \( \bar{\alpha} \) was then applied, plus a random shock drawn from a normal distribution with mean zero and standard deviation \( \bar{\sigma} \). If the resulting state variable lay beyond either bound, an adjustment was made and the state variable was reset at the target. The total number of adjustments for each period was then recorded. Otherwise, this new value became the starting value for the next period. Since wealth is normalized to one, the inverse of the exponential of the target gives the size of purchase relative to wealth.\(^5\) Summing across individual gave aggregate expenditure. In order to eliminate the effect of starting values, this process was repeated until aggregate expenditure has reached the steady state.\(^6\) Then an aggregate shock to wealth was introduced and the resulting dynamics recorded. This process is repeated 5000 times and the average response is reported. This practice can minimize the variation of random sampling.

Each period in the simulation corresponds to a quarter. All parameters are converted to corresponding quarterly values.

---

\(^4\) The standard deviation exceeds that used in other research, probably reflecting the young age of the households in the sample.

\(^5\) For simplicity, I did not consider trade-in values of used vehicles.

\(^6\) It took about 20-40 quarters depending on the underlying parameters of the model.
Before presenting the main results, I will briefly discuss aggregate dynamics implied by other models in the literature for comparison.

6.2 Model without Adjustment Costs

Expenditure on the stock of consumer durables consists of two terms:

\[ C_{t+1} = \delta K_t + (K_{t+1} - K_t) \]  \hspace{1cm} (6.1)

where \( C \) is expenditure, \( K \) is durable stock, and \( \delta \) is a constant depreciation rate, and \( t \) is a time subscript. The first term \( \delta K_t \) is the depreciation of the stock \( K_t \), the second term is the net addition made to the stock, \( K_{t+1} - K_t \).

Mankiw (1982) showed that if there is no adjustment cost, momentary utility function is quadratic in the stock, the objective function is time-separable, the real interest rate equals the subjective discount rate, and capital markets are perfect, then \( K_{t+1} \) follows a random walk. It then from equation (6.1) follows that \( C_{t+1} \) is an IMA(1,1) process:

\[ \Delta C_{t+1} = \epsilon_{t+1} - (1 - \delta)\epsilon_t \]  \hspace{1cm} (6.2)

where \( \epsilon_{t+1} \) is an independently and identically distributed error term proportional to the increment to permanent income in period \( t + 1 \). That is, \( \epsilon \) should capture the change in desired durable stock due to revisions in wealth (or alternatively
permanent income.) At the steady state, if the target level of $K$ is $K^*$, steady state value of $C$, $C^*$ will be equal to $\delta K^*$. That is, each period the household will only purchase to make up for the depreciated part of durable stock.

Suppose the annual depreciation rate is 10 percent. A shock corresponds to a 1 percentage point increase in target stock. The initial steady state level of expenditure is normalized to zero. Figure 8-a plots the dynamic of expenditure after an aggregate shock. The aggregate expenditure initially increases by 40, and then returns to the new steady state in the second period. After the shock aggregate expenditure will be slightly higher than the previous level. The increase equals the annualized unit increment to wealth, as is predicted by PIH/LCH.

In the steady state the representative agent only purchases enough to replace the depreciated stock of consumer durables. As a result, whenever the target stock is revised upwards, there is a “catch-up” effect. That is, in the period of the shock, not only the depreciated portion has to be replaced, additional purchase also has to be made to push the durable stock to the new target level. When the depreciation rate is low, this effect can be very large. As a result, the increase in expenditure following

---

7 The original format of the model is in terms of level. This is the direct result of using quadratic utility function. If a more flexible form such as CRRA is used, one can derive similar relationship in percentage changes, which is more appealing to empirical analysis. As a result, I will use these formula as percentage changes instead in the following analysis.

8 The solid line is the dynamics of aggregate expenditure, and the dashed line is the steady state level. Please note that neither the horizontal time range nor the vertical scale is the same across all figures.
the shock is mainly due to the "catch-up" effect. Moreover, all the adjustment is
done in one period, and no shock has any effect beyond the period when it occurs,
except for its contribution to the revision of wealth.

The series plotted in figure 7 clearly does not exhibit the violent kind of dynamics
portrayed in figure 8-a. Rather, the observed series is smooth and persistent.

6.3 Model With Convex Adjustment Costs

In Bernanke (1984) model, expenditures on consumer durable goods are a fraction
\( \lambda \) of the gap between the desired stock \( K_{t+1}^* \) and the current stock \( K_t \):

\[
C_{t+1} = \lambda(K_{t+1}^* - K_t)
\]  \hspace{1cm} (6.3)

As in Mankiw's model, the desired stock of consumer durable goods follows a random
walk. Equation (6.1) and (6.3) then imply that expenditure on consumer durables
are the ARIMA(1,1,1) process:

\[
\Delta C_{t+1} = (1 - \delta - \lambda)\Delta C_t + \lambda \epsilon_{t+1} - \lambda(1 - \delta)\epsilon_t
\]  \hspace{1cm} (6.4)

where \( \epsilon \) has the same property as in equation (6.2). According to equation (6.1),
households purchase the depreciation on their holdings of consumer durables. As a
result, there is still a "catch-up" effect in this model. However, instead of taking
place in one period, it is spread over the entire future by the partial adjustment.
For this reason, the initial response is much smaller than the previous model.
Bernanke (1984) estimates $\lambda$ to be close to 0.7. Using this parameter value, figure 8-b plots the response to an aggregate shock in target stock. The initial response drops from 40 to 26. This is a significant reduction, but still it is too large compared with the observed values. On the pro side, this model can generate longer dynamics after the shock. This is because with partial adjustment the "catch-up" is only accomplished in several steps.

6.4 Heterogeneous Model

Using the parameters estimated in the previous chapter, the aggregate dynamics for the heterogeneous model can be simulated along the line described at the beginning of this chapter. Figures 8-c and 8-d plot the dynamics after the shock for $\gamma = 0.74$ and $\gamma = 1.37$, respectively.

These dynamics are very different from those shown in the previous two subsections. First of all, the initial response to the shocks is much smaller than before and closer to the observed value. When $\gamma$ is equal to 0.74, aggregate expenditure can be 1.6 percent larger than before in the first period. This implies an elasticity of 1.6. For $\gamma$ equal to 1.37, this elasticity is 4. By construction, these elasticity should be due to changes in the number of adjustments. That is to say, even though there is very little change in target level (due to little change in revision in wealth), it still makes some households to advance their purchases. Second, shocks have long
lasting effect on aggregate expenditure. Unlike the partial adjustment model, this is caused by changes in the number of adjustments due to the shocks. Since households no longer make marginal purchases, the "catch-up" effect does not exist in this model. In terms of persistence of shocks, it takes about 2-4 years to absorb the unit aggregate shock.

The cases $\gamma = 0.74$ and $\gamma = 1.37$ differ in that the former has larger idiosyncratic uncertainty than the latter. As Bertola and Caballero (1990) has shown, larger idiosyncratic shocks result in faster convergence. This fact is also evident in figure 8, where $8d$ exhibits slower convergence than $8c$.

In summary, the heterogeneous model starts from a solid micro background and is capable of generating aggregate dynamics similar to those observed in aggregate data. However, there are several problems that may make it difficult, if not impossible, to compare the overall goodness of fit of the model with the empirical data. First of all, the sample used in the estimation in the previous chapter is not representative of the total population. Second, cross section variation in income is not included. Third, the aggregate data only include expenditure on new motor vehicles and parts, while the model, as well as the micro panel data used to estimate the

---

9 See data appendix C for more detail.
model, also include expenditure on used motor vehicles. Nevertheless, the heterogeneous model does appear to generate aggregate dynamics qualitatively and quantitatively similar to those observed in the data.

6.5 Homogeneous Model

As a final comparison, the homogeneous model is also simulated in the same manner as the heterogeneous model. The result is presented in figures 8-e and 8-f. In addition, for both the heterogeneous and the homogeneous model, the steady state proportion of population who adjust is reported in table 12.

The homogeneous model can be viewed as the constrained version of the heterogeneous model, in which every household must have the same long run distribution. As a result, the (S,s) band will be much wider and household will adjust less often. This is evident in the proportion of households who adjust in steady state, shown in Table 7. When \( \gamma = 0.74 \), only 6.0 percent households adjust each year, which is too small compared with reports from relevant research. For the same reason, the elasticity of aggregate expenditure is larger than for the heterogeneous model. Eberly (1994) also showed that the homogeneous model is capable of generating desirable aggregate dynamics. However, inferiority of the model can be demonstrated in at least two aspects. First, the individual observed targets and band widths show significant heterogeneity, as in figure 5. In particular, individual band width is much
smaller than the support of homogeneous ergodic distribution. Second, the homogeneous model yields too few adjustments per period in the steady state, indicating the uniform bandwidth is too large.
Chapter 7

CONCLUSION AND EXTENSION

It has long been realized that in the presence of non convex adjustment costs, households may adopt the \((S,s)\) rule for durable goods consumption. But so far little work has been done to recover the effect of both observable and unobservable heterogeneity on this choice. On the other hand, the latest literature has introduced a brand new method, which enables us to look at the issue from another angle.

In this thesis I first re-evaluated the optimal decision rules implied by the theoretical model. One major implication of this model is that higher adjustment cost prompt wider inaction range, \textit{ceteris paribus}. Then two methods, the threshold autoregression method and the mixture distribution method, are used in a panel data set to evaluate the validity of the theoretical model. Even though the first method has found some evidence for the model, the more elaborate mixture distribution has produced richer results. This work is the first to use the ergodic distribution
derived from a controlled diffusion process to infer target and band width choices. The model incorporates both observable and unobservable heterogeneity into the determination of the target and band width. I find that observed heterogeneity affect (S,s) rule in a manner consistent with the existing theory. Both forms of heterogeneity prove to be important in explaining households' (S,s) rules. However, it seems that most of the heterogeneity remain unobserved.¹ Moreover, the simulation based on the estimated model has also achieved a much better fit of the observed data.

The estimated model is also capable of reconciling the difference between individual purchasing behavior and aggregate expenditure dynamics. A long standing puzzle is that when aggregate expenditure data are used to test representative-agent models of durable goods consumption, the persistence is much greater than is observed in the micro data. That is to say, if the aggregate behavior can be treated as the same as micro representative agent behavior, this micro agent would have to be adjusting durable stocks continually. In addition, the elasticity of expenditure to income shocks is too large to be justified by PIH/LCH. However, almost all micro panel data sets have shown that households actually adjust only every couple of years. As a result, macro dynamics are different from individual micro dynamics.

Starting from individual households following (S,s) rules (which is consistent with

¹ Large unexplained variance in the model may reflect large measurement errors in ŷ as well as large unobservable heterogeneity in households. These two cannot be distinguished within the current framework.
a modified version of PIH/LCH), this model is capable of generating aggregate dynamics similar to those observed in the data. In particular, it has pinned down the cause of the "excessive" sensitivity to changes in number of adjustments instead of average expenditure per purchase. That is to say, over-sensitivity is just an aggregation phenomenon, and it does not reflect true micro behavior. Finally, the same changes can also explain the long lasting effect of shocks. The persistence of the shock will in general depend on the size of idiosyncratic shocks.

This thesis has developed a new approach to estimate (S,s) models that can be further explored in several other fields. (S,s) model is widely used in stochastic control problems such as optimal inventory policy and optimal investment policy. If the underlying stochastic process can be approximated by the simple Ito process, then the same method can be used to recover the boundary conditions. One advantage of this model lies in its flexibility with data. For example, firm data on capital holdings would be interesting to examine within this framework. It would also be interesting to compare the results from this model with those from the traditional Tobit models.
APPENDICES

Appendix A: Numerical Simulations

This part follows Grossman and Laroque (1990).

Searching for the triple \((y_1, y^*, y_2)\) can be carried out using the following system of differential equations and boundary conditions. Repeat from chapter 2:

\[
-\frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 \left( \frac{h'(y)}{h''(y)} \right)^2 + r(y + \lambda - 1)h'(y) - \bar{\theta}h(y) + \frac{1}{a} = 0 \quad (7.1)
\]

for \(y \in [y_1, y_2], \ y \neq 1 - \lambda; \)

\(h(y) \geq My^a \) for all \(y; \) \hspace{1cm} (7.2)

\(h(y_i) = M y_i^a \) for \(i = 1, 2; \) \hspace{1cm} (7.3)

\(h'(y_i) = aM y_i^{a-1} \) for \(i = 1, 2; \) \hspace{1cm} (7.4)

and \(M\) must satisfy:

\[M = (y^* + \lambda)^{-a} h(y^*) = \sup_y (y + \lambda)^{-a} h(y) \quad (7.5)\]

First, \(M\) can be shown to be bounded inside the following range:

\[\nu \geq M \geq \nu_2 \quad (7.6)\]

where

\[
\nu = \frac{1}{a(\delta + \tau_f)} \left( \frac{(r_f + \delta)(1-a)}{\beta} \right)^{1-a} \quad (7.7)
\]

\[
\nu_2 = \frac{1}{\sigma} = \frac{1}{\sigma + a\delta} \quad \text{if} \ a > 0 \quad (7.8)
\]

\[
\nu_2 = \frac{(1 - e^{\delta/2a})^a}{1 - e^{-\delta/a}} \frac{1 - e^{-\delta}}{\bar{\theta}} \quad \text{if} \ a < 0 \quad (7.9)
\]
Within this range, pick any $M$ (it turns out in most of the cases $M$ is fairly close to the upper bound $v$). Then pick any $y$ as starting value $y_1$. Value matching condition (7.3) defines $h(y_1)$, and smoothing pasting condition (7.4) defines $h'(y_1)$. With $y_1$, $h(y_1)$ and $h'(y_1)$ “known”, substituting them into equation (7.1) can obtain $h''(y_1)$. Pick another $y$ slightly larger than $y_1$ (e.g., I use 0.01 as default increment). Let’s call it $y'$; then $h(y')$, $h'(y')$ can be approximated using 2nd order and 1st order Taylor’s expansion around $y_1$ respectively. Then $h''(y')$ can be found by using equation (7.1). Next, add 0.01 to $y'$ to make $y''$, and repeat the procedure. If at any point $y''$ satisfies equations (7.3) and (7.4), then we are done. $y''$ is the upper bound $y_2$. If such a $y''$ does not exist, change starting value $y_1$ and do over again. If still no such $y''$ exists, change to another $M$; sometimes it seems some $y''$ is a candidate solution, but $M^*$ calculated from equation (7.5) is very different from the assumed $M$, then this $y''$ is also not a solution and new $M$ should be tried. So, after the right $y_1$ and $y_2$ are found, assumed $M$ should be very close to $M^*$ and $y^*$ can be solved using equation (7.5). Proofs of existence and uniqueness of a solution can be found in the original paper.

GAUSS is used to do the numerical search. Large amounts of simulation results are reported in the original paper for various ranges of parameter values. To check the correctness and accuracy of my program, all threshold results are reproduced and all match their results with high accuracy (mostly ±0.05). Note there is only one overlap between table 1 and those in the original paper. The original paper does not produce many results for realistic parameter ranges for this paper’s purpose.

Another slightly easier numerical simulation involve equation (2.25) which is repeated below:

$$-1 = V'(y)E[dy] + \frac{1}{2}V''(y)Var[dy] \quad (7.10)$$

From the first part of simulation I obtain sequence of $y$, $h(y)$, $h'(y)$, and $h''(y)$. Then $x(y)$ can be calculated by equation (2.18). Using $Var[dy] = x^2\sigma^2$ and $E[dy] = r(y + \lambda - 1) + xu$, as well as boundary conditions $V(y_1) = V(y_2) = 0$, equation (7.10) can be solved as follows: First, we already know $y_1$ and $V(y_1)$, so guess a $V'(y_1)$. Substitute this guess into equation (7.10) and obtain $V''(y_1)$. Next, $y''$ is just $(y_1 + 0.01)$ in my simulation (finer increments have been tried with little marginal gain in results but much larger time costs), and $V(y')$ and $V'(y')$ can be approximated using 2nd order and 1st order Taylor series expansion respectively just like before. Finally, $V''(y')$ is found by using equation (7.10) again. Continue this until upper bound $y_2$ is reached. If it happens that $V(y_2) = 0$, then we are done; otherwise, revise the initial guess on $V'(y_1)$. This part is simulated using GAUSS OPTMUM routine. Default specifications (BFGS optimization method with STEPBT line search) are used unmodified since convergence seems fast and accurate. Only one original result is tried to be reproduced due to time constraint.
at this point, and the match is very close (5.73:5.79). All other results reported in table 1 show the right magnitude and pattern. As a result, I am confident of the correctness and accuracy of these results.
Appendix B: Ergodic Distribution

This part has found reference in Dixit and Pindyck (1994), Bertola and Caballero (1990) and Eberly (1994).

I intend to derive the ergodic distribution of $\tilde{y}$, which follows a controlled diffusion process. Following Dixit and Pindyck (1994) and Bertola and Caballero (1990), we can think of the diffusion process as the limit of a random walk with drift. Then the Markov property of this process allows us to define a state transition matrix that governs the dynamics of the state changes. As long as this matrix is regularly recursive, which approximately means that any state in the state space can be reached with positive probability from any starting state, there exists an invariant distribution on the state space. For each individual, this distribution is ergodic in the sense that it is the long run stationary distribution. In the special case without aggregate shocks, it is also the stationary cross-section distribution. However, when there are aggregate shocks, even though each individual still converges to its ergodic distributions, they are constantly reshaped by the aggregate shocks. As a result, the cross-section distribution is not invariant over time.

Rewrite the Ito process for $\tilde{y}$:

$$d\tilde{y} = \tilde{\alpha} dt + \tilde{\sigma} dz$$

(7.11)

where $\tilde{\alpha}$ and $\tilde{\sigma}$ is instantaneous drift and standard deviation, respectively. $dz$ is the increment of standard Brownian motion. First, I define the appropriate random walk discretization of the Ito process:

$$dh = \tilde{\sigma} \sqrt{dt}$$

(7.12)

where $dh$ and $dt$ are infinitesimal step length and time elapse. And

$$\tilde{y}_{t+dt} = \tilde{y}_t + dh \text{ with probability } p = \frac{1}{2} \left( 1 + \frac{\tilde{\sigma}}{\tilde{\sigma}} \sqrt{dt} \right)$$

(7.13)

$$\tilde{y}_{t+dt} = \tilde{y}_t - dh \text{ with probability } q = \frac{1}{2} \left( 1 - \frac{\tilde{\sigma}}{\tilde{\sigma}} \sqrt{dt} \right)$$

(7.14)

The specification in equation (19) and (20) can guarantee convergence to equation (17) in the limit (when both $dt$ and $dh$ go to zero). Intuitively, at any point of time
can either jump up by $dh$, or jump down by $dh$, with different probabilities $p$ and $q$ respectively. This binomial distribution converges to an Ito's process as the step length $dh$ and the time length $dt$ approaches zero.

First, consider the process specified in equations (19) and (20) in general. Let $\phi(\tilde{y})$ be the density function of $\tilde{y}$; then in the neighborhood of any $y$ it should follow:

$$\phi(\tilde{y}) = p \cdot \phi(\tilde{y} - dh) + q \cdot \phi(\tilde{y} + dh)$$  \hspace{1cm} (7.15)

because any state $\tilde{y}$ can be reached either from state $\tilde{y} - dh$ with probability $p$, or from state $\tilde{y} + dh$ with probability $q$. Substitute in $p$ and $q$ and expand the right-hand side with Taylor's expansion. After collecting terms, we obtain:

$$\phi(\tilde{y}) = \frac{1}{2} [1 + \frac{\tilde{\alpha}}{\sigma^2} dh] [\phi(\tilde{y}) - \phi'(\tilde{y}) dh + \frac{1}{2} \phi''(\tilde{y}) (dh)^2 + \ldots]$$

$$+ \frac{1}{2} [1 - \frac{-\tilde{\alpha}}{\sigma^2} dh] [\phi(\tilde{y}) + \phi'(\tilde{y}) dh + \frac{1}{2} \phi''(\tilde{y}) (dh)^2 + \ldots]$$

$$= \phi(\tilde{y}) - \frac{\tilde{\alpha}}{\sigma^2} \phi'(\tilde{y}) (dh)^2 + \frac{1}{2} \phi''(\tilde{y}) (dh)^2 + \ldots$$  \hspace{1cm} (7.16)

The terms indicated by $\ldots$ all go to zero faster than $(dh)^2$. Cancel out $\phi(\tilde{y})$ on both sides and take limits as $dh \to 0$. This gives the following differential equation:

$$\phi''(\tilde{y}) = \gamma \phi'(\tilde{y})$$  \hspace{1cm} (7.17)

where $\gamma = \frac{-2\tilde{\alpha}}{\sigma^2}$.²

The general solution to this type of differential equations is:

$$\phi(\tilde{y}) = A \exp(\gamma \tilde{y}) + B$$  \hspace{1cm} (7.18)

where $A$ and $B$ are constants to be determined. We need to use the boundary conditions to pin down them. First consider the lower bound $\tilde{y}^l$. It can only be reached by jumping down from $\tilde{y}^l + dh$, so:

$$\phi(\tilde{y}^l) = q \cdot \phi(\tilde{y}^l + dh)$$  \hspace{1cm} (7.19)

Using Taylor's expansion and rearrange, then take limit as $dh \to 0$. It is easy to verify that:

$$\phi(\tilde{y}^l) = 0$$  \hspace{1cm} (7.20)

Similarly, we can show at the upper bound $\tilde{y}^u$:

$$\phi(\tilde{y}^u) = 0$$  \hspace{1cm} (7.21)

²Same result can be obtained by directly using Kolmogorov Equations. See Dixit and Pindyck (1994), page 88-92.
In addition, at the internal return point \( \bar{y}^* \), not only it can be reached from its own left or right, but also from either bound after the state variable hits. So it must be true that:

\[
\phi(\bar{y}^*) = p \cdot \phi(\bar{y}^* - dh) + q \cdot \phi(\bar{y}^* + dh) + p \cdot \phi(\bar{y}^u - dh) + q \cdot \phi(\bar{y}^d + dh)
\]  

(7.22)

Using the same technique as before, it is easy to verify:

\[
\phi'_-(\bar{y}^*) = \phi'_+ (\bar{y}^*) + \phi'_+(\bar{y}^d) - \phi'_-(\bar{y}^u)
\]  

(7.23)

where a "-" or "+" sign indicates left or right derivative respectively.

Since density function can never go below zero, and by equation (26) and (27) \( \phi'_+(\bar{y}^d) \) must be positive and \( \phi'_-(\bar{y}^u) \) be negative. As a result, \( \phi'_-(\bar{y}^*) \) is different from \( \phi'_+(\bar{y}^*) \). That is to say, \( \phi(\bar{y}) \) function is not differentiable at the internal return point \( \bar{y}^* \). So we break equation (24) into two pieces and derive the parameters separately.

\[
\phi(\bar{y}) = A_l \exp(\gamma \bar{y}) + B_l \quad \text{if} \ \bar{y} \in (\bar{y}^l, \bar{y}^r)
\]

(7.24)

\[
\phi(\bar{y}) = A_u \exp(\gamma \bar{y}) + B_u \quad \text{if} \ \bar{y} \in (\bar{y}^r, \bar{y}^u)
\]

(7.25)

Finally, we can impose the condition that the density should sum up to unity:

\[
\int_{-\infty}^{+\infty} \phi(\bar{y}) d\bar{y} = 1
\]  

(7.26)

which can be rewritten as:

\[
\int_{\bar{y}^l}^{\bar{y}^u} \phi(\bar{y}) d\bar{y} = \int_{\bar{y}^l}^{\bar{y}^r} (A_l \exp(\gamma \bar{y}) + B_l) d\bar{y} + \int_{\bar{y}^r}^{\bar{y}^u} (A_u \exp(\gamma \bar{y}) + B_u) d\bar{y} = 1
\]  

(7.27)

The system of equation (26), (27), (29) and (33) turns out to be sufficient to pin down the underlying parameters of the density function. Specifically, to simplify the notation, let:

\[
l = \exp(\gamma \bar{y}^l), \quad s = \exp(\gamma \bar{y}^r), \quad u = \exp(\gamma \bar{y}^u)
\]  

(7.28)

then the parameters can be solved as:

\[
A_l = \frac{u - s}{u(s - l)(\bar{y}^u - \bar{y}^r) - l(u - s)(\bar{y}^r - \bar{y}^l)}
\]

(7.29)

\[
B_l = A_l(-l)
\]

(7.30)

\[
A_u = -A_l\left(\frac{s - l}{u - s}\right)
\]

(7.31)

\[
B_u = A_u(-u)
\]

(7.32)
In order to figure out the exact shape of distribution, it is useful to derive the signs of $A_l$ and $A_u$. Obviously the shape depends on the sign of $\gamma$. If $\gamma$ is positive, we can show that $A_l$ is positive, which directly implies that $A_u$ is negative, so the density function increases convexly on the left side, and decreases concavely on the right. On the other hand, if $\gamma$ is negative, then $A_l$ is negative and $A_u$ is positive, so the density function increases concavely on the left, and decreases convexly on the right. For the special case in which $\gamma$ is zero, it is easy to show that the density function is triangular. On the other extreme, if $\gamma$ is very large in absolute value, the ergodic distribution converges to a rectangular distribution (uniform), and the support is either the whole lower band (if $\gamma \to -\infty$) or the whole upper band (if $\gamma \to +\infty$). Here we only prove the first regular case of $\gamma$ being positive, the other proofs are similar and/or straightforward.

If $\gamma$ is positive, equation (34) is monotonic in $\tilde{y}$. As a result, $\tilde{y}'' > \tilde{y}' > \tilde{y}' > u > s > l > 0$. Note that the following must hold for any $x$ as long as it is positive:

$$\ln(x) \geq x - 1$$

$$\ln\left(\frac{1}{x}\right) \leq 1 - x$$

Let $x_1 = \frac{u}{s}$ in equation (39) and $x_2 = \frac{l}{s}$ in equation (40). By the definition of $l$, $s$ and $u$, we know $x_1 > 1$ and $\frac{1}{x_2} > 1$. We can obtain:

$$\ln\left(\frac{u}{s}\right) \geq \frac{u}{s} - 1$$

$$\ln\left(\frac{s}{l}\right) \leq 1 - \frac{l}{s}$$

Note that all four parts of the above equations should be positive, so combining equations (41) and (42) we have:

$$(1 - \frac{l}{s})(\ln\left(\frac{u}{s}\right)) > \ln\left(\frac{s}{l}\right)(\frac{u}{s} - 1)$$

Also, since $u > l > 0$, we can multiply the LHS with $u$ and RHS with $l$, and rewrite the equation as:

$$u(1 - \frac{l}{s})(\ln(u) - \ln(s)) > l(\ln(s) - \ln(l))(\frac{u}{s} - 1)$$

It should be clear by now that equation (44) implies $A_l > 0$.

In summary, the ergodic distribution of $\tilde{y}$ is piece-wise exponential, but non-differentiable at the internal return point $\tilde{y}'$. It starts from the lower bound with density zero, increases to reach the peak at the internal return point, and then decreases to zero as it reaches the upper bound. This shape of $\phi(\tilde{y})$ is the same as those shown in Bertola and Caballero (1990). See figure 4 for examples of this type of distribution.
Appendix C: Data

The data used here comes from National Longitudinal Survey of Youth 1979 Cohorts (NLSY79). Details about the survey and variable definitions can be found in *NLSY79 User’s Guide* prepared by Center for Human Resource Research (CHRR) at the Ohio State University.

The sub-sample I used here includes 6111 youths designed to be representative of the non-institutionalized civilian segment of young people living in the United States in 1979 and born between January 1, 1957, and December 31, 1964. They are interviewed annually from 1979 to 1994, and then once every two years afterwards.

Main series used are described as follows:

**Income**: Respondents are asked to identify their income sources and amounts for themselves and their spouses/partners every year. From these responses, the CHRR staff creates a new variable entitled “Net Family Income”. Before 1987, some respondents who were still living with their parents were given different and shorter income question forms while their parents were asked to answer most of the income questions. After 1987 all respondents are given the same form. Partner’s income is not included in the calculation of Net Family Income. Taking this created variable, I first deduct an estimate of federal income tax from it, calculated using the rates provided in *Individual Income Tax Return 1995, Statistics of Income: Publication 1304*. It reports average federal income tax rates for each income level from 1987 to 1994. Then the after tax incomes are deflated by CPI-U, which is reported by the Bureau of Labor Statistics (Website, 1982-1984=100). This is named “real after tax net family income” in the paper, or default income. Another term, lifetime income, is constructed by taking the present value of a flow equal to current income until age 66 and discounted at an annual rate of 5 percent. Note there is no income data for 1994 since no survey was fielded in 1995. Instead, income in 1994 is constructed by simply averaging the 1993 and 1995 income values.

**Real Net Worth**: Starting from 1985, the survey asks detailed questions about asset holdings, including moving vehicles. Net worth is calculated by subtracting all liabilities from total assets based on the respondents’ reports of their current market values. Then I deflate net worth by CPI-U to obtain real net worth. Note no asset questions were asked in 1991, so real net worth in 1991 is approximated simply by averaging the 1990 and 1992 values.
**Moving Vehicle Stock:** This is actually part of the asset questions. Respondents are asked about current market value of the moving vehicles that they and their spouses/partners own. Note only the extended version of the asset section (which only starts from 1987) provides enough information in this respect, this is one of the major reasons that I limit the time period to be between 1987 and 1994. Then this value is deflated by the CPI for transportation, and it is called the “real vehicle stock”.

**Demographic Variables:** One big advantage of using this data set is its enormous amount of detailed data on demographics. I have chosen to include the following:

- respondent’s age in each year (age).
- respondent’s gender (sex).
- respondent’s race (race: Hispanic, African, or neither).
- highest level of education achieved by the respondent (education).
- currently married or not (marital status).
- total number of adults in the household (18 and older).
- total number of children in the household (17 and younger).
- moving vehicle license registration fee in residence state.

**Lifetime Wealth:** This is constructed by summing up real net worth and lifetime income for each year. All respondents with any top-coded values are omitted.

**Construction of vehicle stock in 1991:** First, compare demographically adjusted real vehicle stock in 1990 and 1992. If the change is dramatic, I assume there is an adjustment either in 1991 or 1992. Second, pooled regression of real vehicle stock on demographic variables and lifetime wealth provides predicted value of real vehicle stock. Then an estimated change in real vehicle stock from 1990 to 1991 can be used to pin down whether the change should occur in 1991 or 1992. After these two steps, it no adjustment is warranted, real vehicle stock is taken to be the average of 1990 and 1992 values. If there is an adjustment in 1991, then real vehicle stock in 1991 is approximated with 1.05 times 1992 value. If there is an adjustment in 1992 instead, real vehicle stock in 1991 is approximated with 0.95 times 1990 value. Here I assume away the probability that back-to-back adjustment has occurred because it is very rare and inclusion of this may complicate the construction. Critical questions remain on how to choose the right cutoffs. I’ve tried a lot of possible combinations to make the cross-section distribution as close to 1990 and 1992 as possible. For example, I try to make the mean, median, skewness and kurtosis as close as possible
to those in 1990 and 1992 cross-section. The result is, if the percentage change from 1990 to 1992 is greater than 30% or less than -40% then I assume there is an adjustment in 1991 or 1992. And if the predicted change in 1991 is positive, then I assume the change happens in 1991; otherwise, it happens in 1992.

Since the vehicle stock is the most important variable in the paper, any artificial construction of data should be carried out with extreme caution. As a result, I do not attempt to construct other missing values other than 1991 and just drop them instead. Since this is mainly a labor survey, it is very common that some individual has missed interview in some years. As a result, only about a quarter of the sample is left with complete data series for 1987-1994. But this still makes the sample size around 1500. To check whether this create a sample selection problem, I check the age, sex, race distribution as well as average income trend, net worth trend and real vehicle stock holding trend. All of the above variables are not significantly different from the sample of 6111 respondents. As a result, sample selection is not a critical issue here.
Appendix D: Tables

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>((y_1, y_*, y_2))</th>
<th>(y_2 - y_1)</th>
<th>(E[T(y^*)])</th>
<th>((\hat{y}_1)^*)</th>
<th>((\hat{y}_2)^*)</th>
<th>((\hat{y}_3)^*)</th>
<th>drift in income</th>
<th>s.d. in income</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda = 0.10 )</td>
<td>((10,40,19,05,33.97))</td>
<td>23.37</td>
<td>3.05</td>
<td>5.25</td>
<td>-39.91</td>
<td>134.66</td>
<td>0.094</td>
<td>0.313</td>
</tr>
<tr>
<td>0.9</td>
<td>((4.89,8.06,13.55))</td>
<td>8.75</td>
<td>3.20</td>
<td>12.41</td>
<td>-21.15</td>
<td>109.13</td>
<td>0.084</td>
<td>0.274</td>
</tr>
<tr>
<td>1.1</td>
<td>((10.03,6.33,10.12))</td>
<td>6.09</td>
<td>3.21</td>
<td>15.80</td>
<td>-7.65</td>
<td>91.85</td>
<td>0.074</td>
<td>0.240</td>
</tr>
<tr>
<td>1.3</td>
<td>((5.82,5.72,8.85))</td>
<td>5.01</td>
<td>3.20</td>
<td>17.48</td>
<td>3.17</td>
<td>79.34</td>
<td>0.067</td>
<td>0.213</td>
</tr>
<tr>
<td>1.5</td>
<td>((3.79,5.46,6.21))</td>
<td>4.42</td>
<td>3.17</td>
<td>18.32</td>
<td>11.82</td>
<td>69.87</td>
<td>0.060</td>
<td>0.188</td>
</tr>
<tr>
<td>1.7</td>
<td>((3.83,5.36,6.79))</td>
<td>4.06</td>
<td>3.12</td>
<td>18.65</td>
<td>18.89</td>
<td>62.46</td>
<td>0.055</td>
<td>0.169</td>
</tr>
<tr>
<td>2.1</td>
<td>((3.91,5.34,7.71))</td>
<td>3.80</td>
<td>3.06</td>
<td>18.73</td>
<td>24.77</td>
<td>56.50</td>
<td>0.051</td>
<td>0.153</td>
</tr>
<tr>
<td>2.3</td>
<td>((4.00,5.35,7.63))</td>
<td>3.63</td>
<td>3.00</td>
<td>18.69</td>
<td>25.73</td>
<td>51.58</td>
<td>0.047</td>
<td>0.140</td>
</tr>
<tr>
<td>2.5</td>
<td>((4.10,5.40,7.49))</td>
<td>3.51</td>
<td>2.94</td>
<td>18.52</td>
<td>34.00</td>
<td>47.48</td>
<td>0.044</td>
<td>0.128</td>
</tr>
<tr>
<td>2.7</td>
<td>((4.20,5.45,7.62))</td>
<td>3.42</td>
<td>2.91</td>
<td>18.35</td>
<td>37.68</td>
<td>43.98</td>
<td>0.042</td>
<td>0.118</td>
</tr>
<tr>
<td>2.9</td>
<td>((4.30,5.52,7.59))</td>
<td>3.35</td>
<td>2.86</td>
<td>18.12</td>
<td>40.91</td>
<td>40.97</td>
<td>0.040</td>
<td>0.110</td>
</tr>
<tr>
<td>3.1</td>
<td>((4.40,5.59,7.70))</td>
<td>3.30</td>
<td>2.83</td>
<td>17.89</td>
<td>43.76</td>
<td>38.35</td>
<td>0.038</td>
<td>0.103</td>
</tr>
<tr>
<td>( \lambda = 0.05 )</td>
<td>((6.90,18.21,43.98))</td>
<td>37.08</td>
<td>5.70</td>
<td>5.49</td>
<td>-39.31</td>
<td>133.82</td>
<td>0.094</td>
<td>0.312</td>
</tr>
<tr>
<td>0.9</td>
<td>((3.30,7.57,17.06))</td>
<td>13.76</td>
<td>5.79</td>
<td>13.21</td>
<td>-20.49</td>
<td>107.28</td>
<td>0.083</td>
<td>0.272</td>
</tr>
<tr>
<td>1.1</td>
<td>((2.55,9.00,12.44))</td>
<td>9.59</td>
<td>5.64</td>
<td>16.95</td>
<td>-6.71</td>
<td>89.77</td>
<td>0.074</td>
<td>0.238</td>
</tr>
<tr>
<td>1.3</td>
<td>((2.76,3.10,14.66))</td>
<td>7.90</td>
<td>5.48</td>
<td>18.83</td>
<td>3.85</td>
<td>77.32</td>
<td>0.066</td>
<td>0.210</td>
</tr>
<tr>
<td>1.5</td>
<td>((2.77,5.05,8.83))</td>
<td>7.06</td>
<td>5.39</td>
<td>19.80</td>
<td>12.23</td>
<td>67.97</td>
<td>0.060</td>
<td>0.186</td>
</tr>
<tr>
<td>1.7</td>
<td>((2.83,9.34,9.57))</td>
<td>6.34</td>
<td>5.29</td>
<td>20.24</td>
<td>19.05</td>
<td>60.71</td>
<td>0.055</td>
<td>0.167</td>
</tr>
<tr>
<td>1.9</td>
<td>((2.91,9.23,9.12))</td>
<td>5.91</td>
<td>5.18</td>
<td>20.33</td>
<td>24.76</td>
<td>54.92</td>
<td>0.051</td>
<td>0.152</td>
</tr>
<tr>
<td>2.1</td>
<td>((3.00,9.44,8.98))</td>
<td>5.58</td>
<td>5.08</td>
<td>20.41</td>
<td>25.60</td>
<td>50.16</td>
<td>0.047</td>
<td>0.138</td>
</tr>
<tr>
<td>2.3</td>
<td>((3.10,9.49,8.91))</td>
<td>5.11</td>
<td>4.98</td>
<td>20.04</td>
<td>33.78</td>
<td>46.18</td>
<td>0.044</td>
<td>0.127</td>
</tr>
<tr>
<td>2.5</td>
<td>((3.19,9.05,8.91))</td>
<td>4.72</td>
<td>4.92</td>
<td>19.80</td>
<td>37.42</td>
<td>42.78</td>
<td>0.041</td>
<td>0.117</td>
</tr>
<tr>
<td>2.7</td>
<td>((3.29,5.13,8.02))</td>
<td>4.63</td>
<td>4.83</td>
<td>19.49</td>
<td>40.64</td>
<td>39.87</td>
<td>0.039</td>
<td>0.109</td>
</tr>
<tr>
<td>2.9</td>
<td>((3.40,5.59,7.70))</td>
<td>4.59</td>
<td>4.78</td>
<td>19.19</td>
<td>43.49</td>
<td>37.32</td>
<td>0.037</td>
<td>0.102</td>
</tr>
<tr>
<td>( \lambda = 0.10 )</td>
<td>((5.10,17.37,51.45))</td>
<td>46.45</td>
<td>7.40</td>
<td>5.75</td>
<td>-39.02</td>
<td>133.26</td>
<td>0.093</td>
<td>0.311</td>
</tr>
<tr>
<td>0.9</td>
<td>((2.54,7.24,19.41))</td>
<td>16.87</td>
<td>7.23</td>
<td>13.81</td>
<td>-20.49</td>
<td>106.22</td>
<td>0.083</td>
<td>0.271</td>
</tr>
<tr>
<td>1.1</td>
<td>((2.22,5.00,14.07))</td>
<td>11.85</td>
<td>7.07</td>
<td>17.88</td>
<td>-6.34</td>
<td>88.48</td>
<td>0.074</td>
<td>0.237</td>
</tr>
<tr>
<td>1.3</td>
<td>((2.17,5.02,12.01))</td>
<td>9.84</td>
<td>6.91</td>
<td>19.92</td>
<td>4.00</td>
<td>75.99</td>
<td>0.066</td>
<td>0.209</td>
</tr>
<tr>
<td>1.5</td>
<td>((2.11,4.79,10.96))</td>
<td>8.75</td>
<td>6.70</td>
<td>20.88</td>
<td>12.30</td>
<td>66.76</td>
<td>0.060</td>
<td>0.186</td>
</tr>
<tr>
<td>1.7</td>
<td>((2.17,4.69,10.43))</td>
<td>8.16</td>
<td>6.59</td>
<td>21.32</td>
<td>19.12</td>
<td>59.55</td>
<td>0.055</td>
<td>0.167</td>
</tr>
<tr>
<td>1.9</td>
<td>((2.35,4.67,10.12))</td>
<td>7.77</td>
<td>6.47</td>
<td>21.41</td>
<td>24.79</td>
<td>53.80</td>
<td>0.050</td>
<td>0.151</td>
</tr>
<tr>
<td>2.1</td>
<td>((2.43,4.69,9.98))</td>
<td>7.55</td>
<td>6.41</td>
<td>21.32</td>
<td>29.62</td>
<td>49.06</td>
<td>0.047</td>
<td>0.137</td>
</tr>
<tr>
<td>2.3</td>
<td>((2.52,4.75,9.89))</td>
<td>7.37</td>
<td>6.28</td>
<td>21.05</td>
<td>33.81</td>
<td>45.13</td>
<td>0.044</td>
<td>0.126</td>
</tr>
<tr>
<td>2.5</td>
<td>((2.61,4.83,9.65))</td>
<td>7.24</td>
<td>6.15</td>
<td>20.70</td>
<td>37.49</td>
<td>41.80</td>
<td>0.041</td>
<td>0.115</td>
</tr>
<tr>
<td>2.7</td>
<td>((2.70,4.91,9.35))</td>
<td>7.15</td>
<td>6.06</td>
<td>20.37</td>
<td>40.72</td>
<td>38.91</td>
<td>0.039</td>
<td>0.107</td>
</tr>
<tr>
<td>2.9</td>
<td>((4.40,5.59,7.70))</td>
<td>7.13</td>
<td>5.99</td>
<td>20.00</td>
<td>43.60</td>
<td>36.40</td>
<td>0.037</td>
<td>0.100</td>
</tr>
</tbody>
</table>

Table 1: Numerical Simulation Results

\((\delta = 0.10)\)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>N</th>
<th>Mean</th>
<th>Std.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>$\ln \frac{w}{K}$</td>
<td>11792</td>
<td>2.13</td>
<td>0.90</td>
<td>-1.06</td>
<td>7.45</td>
</tr>
<tr>
<td>AGE</td>
<td>year</td>
<td>11792</td>
<td>29.83</td>
<td>3.15</td>
<td>23</td>
<td>37</td>
</tr>
<tr>
<td>RACE1</td>
<td>1 if black</td>
<td>11792</td>
<td>0.05</td>
<td>0.22</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RACE2</td>
<td>1 if hispanic</td>
<td>11792</td>
<td>0.07</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SEX</td>
<td>1 if male</td>
<td>11792</td>
<td>0.51</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MARI</td>
<td>1 if married</td>
<td>11788</td>
<td>0.65</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EDUC</td>
<td>1 if college up</td>
<td>11779</td>
<td>0.20</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NA</td>
<td>number of adults</td>
<td>11736</td>
<td>1.95</td>
<td>0.82</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>NC</td>
<td>number of children</td>
<td>11788</td>
<td>1.08</td>
<td>0.15</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>LF</td>
<td>(log) license fee</td>
<td>9976</td>
<td>3.99</td>
<td>0.27</td>
<td>2.87</td>
<td>4.61</td>
</tr>
</tbody>
</table>

*Note: Sample period is 1987-1994 (8 years). Marriage status is current.*

**Table 2: Sample Statistics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>mean</th>
<th>std.</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>2.68</td>
<td>0.93</td>
<td>2.02</td>
<td>2.54</td>
<td>3.21</td>
</tr>
<tr>
<td>$\theta_r$</td>
<td>1.97</td>
<td>0.70</td>
<td>1.50</td>
<td>1.86</td>
<td>2.35</td>
</tr>
<tr>
<td>$t$</td>
<td>1.74</td>
<td>0.71</td>
<td>1.27</td>
<td>1.65</td>
<td>2.12</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.19</td>
<td>0.83</td>
<td>-0.11</td>
<td>0.22</td>
<td>0.51</td>
</tr>
<tr>
<td>$\sigma_r$</td>
<td>0.26</td>
<td>0.35</td>
<td>0</td>
<td>0.14</td>
<td>0.40</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>0.75</td>
<td>0.53</td>
<td>0.34</td>
<td>0.64</td>
<td>1.01</td>
</tr>
</tbody>
</table>

**Table 3: Threshold Estimation Result**

*(in natural logarithm)*
<table>
<thead>
<tr>
<th>Variables</th>
<th>const.</th>
<th>$\varphi_1(alw)$</th>
<th>$\varphi_2(drift)$</th>
<th>$\varphi_3(std.)$</th>
<th>$\varphi_4(lf)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates (std. err.)</td>
<td>0.28 (0.55)</td>
<td>-0.09 (0.11)</td>
<td>0.26 (0.56)</td>
<td>0.42 (0.17)</td>
<td>0.38 (0.15)</td>
</tr>
</tbody>
</table>

*Note:* Heteroskedasticity corrected.

**Table 4: Determinants of Bandwidth**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Bound (cdf=2.8%)</td>
<td>1.43</td>
<td>1.47</td>
<td>1.77</td>
<td>1.96</td>
<td>1.97</td>
<td>1.92</td>
<td>1.67</td>
<td>1.88</td>
</tr>
<tr>
<td>Upper Bound (cdf=97.2%)</td>
<td>85.44</td>
<td>80.45</td>
<td>86.94</td>
<td>63.97</td>
<td>74.12</td>
<td>69.82</td>
<td>74.71</td>
<td>68.39</td>
</tr>
</tbody>
</table>

*Note:* 1991 values calculated using the constructed $y$.

**Table 5: Minimum Requirement on Thresholds**
<table>
<thead>
<tr>
<th>Variable</th>
<th>Target</th>
<th>Bandwidth</th>
<th>Target</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONST.</td>
<td>0.10870</td>
<td>(0.17275)</td>
<td>0.12791</td>
<td>(0.18193)</td>
</tr>
<tr>
<td></td>
<td>-0.15321</td>
<td>(0.26147)</td>
<td>-0.11660</td>
<td>(0.30882)</td>
</tr>
<tr>
<td>AGE</td>
<td>0.00472</td>
<td>(0.00398)</td>
<td>0.00452</td>
<td>(0.00427)</td>
</tr>
<tr>
<td></td>
<td>-0.15321</td>
<td>(0.00618)</td>
<td>-0.11660</td>
<td>(0.00646)</td>
</tr>
<tr>
<td>RACE1</td>
<td>-0.05787</td>
<td>(0.03945)</td>
<td>-0.03097</td>
<td>(0.04082)</td>
</tr>
<tr>
<td></td>
<td>0.00873</td>
<td>(0.06711)</td>
<td>0.03794</td>
<td>(0.06013)</td>
</tr>
<tr>
<td>RACE2</td>
<td>0.05465</td>
<td>(0.03547)</td>
<td>0.07477*</td>
<td>(0.03727)</td>
</tr>
<tr>
<td></td>
<td>-0.01400</td>
<td>(0.05418)</td>
<td>0.01091</td>
<td>(0.05071)</td>
</tr>
<tr>
<td>SEX</td>
<td>-0.00587</td>
<td>(0.01746)</td>
<td>-0.00434</td>
<td>(0.01758)</td>
</tr>
<tr>
<td></td>
<td>0.00560</td>
<td>(0.02545)</td>
<td>0.00112</td>
<td>(0.02811)</td>
</tr>
<tr>
<td>MARI</td>
<td>-0.16399**</td>
<td>(0.02692)</td>
<td>-0.12132**</td>
<td>(0.02517)</td>
</tr>
<tr>
<td></td>
<td>-0.38733**</td>
<td>(0.04256)</td>
<td>-0.24252**</td>
<td>(0.03801)</td>
</tr>
<tr>
<td>EDUC</td>
<td>0.02592</td>
<td>(0.02086)</td>
<td>0.00881</td>
<td>(0.02533)</td>
</tr>
<tr>
<td></td>
<td>-0.05814*</td>
<td>(0.03296)</td>
<td>-0.07440*</td>
<td>(0.03350)</td>
</tr>
<tr>
<td>NA</td>
<td>0.08356**</td>
<td>(0.01813)</td>
<td>0.07200**</td>
<td>(0.01785)</td>
</tr>
<tr>
<td></td>
<td>0.05228*</td>
<td>(0.02860)</td>
<td>0.01541</td>
<td>(0.02846)</td>
</tr>
<tr>
<td>NC</td>
<td>0.04870**</td>
<td>(0.00992)</td>
<td>0.03353**</td>
<td>(0.00967)</td>
</tr>
<tr>
<td></td>
<td>0.11468**</td>
<td>(0.01563)</td>
<td>0.08146**</td>
<td>(0.01431)</td>
</tr>
<tr>
<td>LF</td>
<td>0.08983**</td>
<td>(0.03032)</td>
<td>0.08271**</td>
<td>(0.03185)</td>
</tr>
<tr>
<td></td>
<td>0.09818*</td>
<td>(0.04657)</td>
<td>0.10860**</td>
<td>(0.05206)</td>
</tr>
</tbody>
</table>

| $\sigma_r$ | 0.28753** | (0.00576) | 0.30762** | (0.00616) |
| $\sigma_b$ | 0.45169** | (0.00904) | 0.45841** | (0.00918) |
| $\rho$     | 0.43729** | (0.02290) | 0.34204** | (0.02501) |

| MEAN LOG LIKELIHOOD | -0.69057 | -0.81679 |
| LR TEST (d.f.=9)    | 158.28924 | 115.57196 |

Note: N = 1247. Standard error in parenthesis, except for the last LR test while it is p-value instead. LR test is for setting all parameters of observed heterogeneity equal to zero. (*) is for significance level 5 percent and (**) for 1 percent.

**Table 6: Direct Estimation With Observed Heterogeneity**
<table>
<thead>
<tr>
<th>Variable</th>
<th>Target</th>
<th>Bandwidth</th>
<th>Target</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONST.</td>
<td>0.71212** (0.00836)</td>
<td>-0.23009** (0.01341)</td>
<td>0.68157** (0.00888)</td>
<td>-0.23009** (0.01341)</td>
</tr>
<tr>
<td>$\sigma_r$</td>
<td>0.29524** (0.00591)</td>
<td>0.31384** (0.00628)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_b$</td>
<td>0.47341** (0.00948)</td>
<td>0.47341** (0.00948)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.45519** (0.02245)</td>
<td>0.35523** (0.02473)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAN LOG LIKELIHOOD</td>
<td>-0.75403</td>
<td>-0.86313</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* N=1247. Standard error in parenthesis. (*) is for significance level 5 percent and (**) for 1 percent.

**Table 7: Direct Estimation Without Observed Heterogeneity**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\gamma = 0.74$</th>
<th>$\gamma = 1.37$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>Bandwidth</td>
<td>Target</td>
</tr>
<tr>
<td>CONST.</td>
<td>0.65254 (0.00105)</td>
<td>0.85460 (0.00127)</td>
</tr>
<tr>
<td>MEAN LOG LIKELIHOOD</td>
<td>-1.20761</td>
<td>-1.21474</td>
</tr>
</tbody>
</table>

*Note:* N=9429. Standard error in parenthesis. All test statistics are significant at 1 percent.

**Table 8: Homogeneous Model**

88
<table>
<thead>
<tr>
<th>Variable</th>
<th>Target Bandwidth</th>
<th>Target Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>0.47582**</td>
<td>0.41111**</td>
</tr>
<tr>
<td></td>
<td>(0.00078)</td>
<td>(0.00185)</td>
</tr>
<tr>
<td>age</td>
<td>-0.00027</td>
<td>0.00014</td>
</tr>
<tr>
<td></td>
<td>(0.00290)</td>
<td>(0.01295)</td>
</tr>
<tr>
<td>race1</td>
<td>-0.01308**</td>
<td>-0.01360**</td>
</tr>
<tr>
<td></td>
<td>(0.00085)</td>
<td>(0.00279)</td>
</tr>
<tr>
<td>race2</td>
<td>0.07114**</td>
<td>0.06693**</td>
</tr>
<tr>
<td></td>
<td>(0.00088)</td>
<td>(0.00195)</td>
</tr>
<tr>
<td>sex</td>
<td>-0.00598</td>
<td>-0.00472</td>
</tr>
<tr>
<td></td>
<td>(0.00622)</td>
<td>(0.00774)</td>
</tr>
<tr>
<td>mari</td>
<td>-0.06843**</td>
<td>-0.04368</td>
</tr>
<tr>
<td></td>
<td>(0.00783)</td>
<td>(0.01038)</td>
</tr>
<tr>
<td>educ</td>
<td>0.01134**</td>
<td>0.02053**</td>
</tr>
<tr>
<td></td>
<td>(0.00096)</td>
<td>(0.00453)</td>
</tr>
<tr>
<td>na</td>
<td>0.01183*</td>
<td>0.01037</td>
</tr>
<tr>
<td></td>
<td>(0.00509)</td>
<td>(0.02284)</td>
</tr>
<tr>
<td>nc</td>
<td>0.02054**</td>
<td>0.01741</td>
</tr>
<tr>
<td></td>
<td>(0.00142)</td>
<td>(0.01403)</td>
</tr>
<tr>
<td>lf</td>
<td>0.05031**</td>
<td>0.05015**</td>
</tr>
<tr>
<td></td>
<td>(0.01921)</td>
<td>(0.01699)</td>
</tr>
<tr>
<td>(\sigma_r)</td>
<td>0.21281**</td>
<td>0.22354***</td>
</tr>
<tr>
<td></td>
<td>(0.00673)</td>
<td>(0.00370)</td>
</tr>
<tr>
<td>(\sigma_b)</td>
<td>0.31239***</td>
<td>0.32640***</td>
</tr>
<tr>
<td></td>
<td>(0.00267)</td>
<td>(0.00280)</td>
</tr>
<tr>
<td>mean log likelihood</td>
<td>-1.07333</td>
<td>-1.07556</td>
</tr>
<tr>
<td>lr test</td>
<td>254.58300</td>
<td>252.88578*</td>
</tr>
<tr>
<td></td>
<td>(d.f.=9)</td>
<td>(d.f.=9)</td>
</tr>
</tbody>
</table>

Note: N=9429. Standard error in parenthesis, except for the last LR test while it is p-value instead. LR test is for setting all parameters of observed heterogeneity equal to zero. (*) is for significance level 5 percent and (**) for 1 percent.

**Table 9: Maximum Likelihood Estimation Results**

**Heterogeneous Model**

89
<table>
<thead>
<tr>
<th>Variable</th>
<th>Target</th>
<th>Bandwidth</th>
<th>Target</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONST.</td>
<td>0.67010</td>
<td>0.23908</td>
<td>0.63202</td>
<td>0.18440</td>
</tr>
<tr>
<td></td>
<td>(0.00344)</td>
<td>(0.00872)</td>
<td>(0.00860)</td>
<td>(0.02136)</td>
</tr>
<tr>
<td>$\sigma_r$</td>
<td>0.22186</td>
<td>0.23092</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00477)</td>
<td>(0.01973)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_b$</td>
<td>0.32993</td>
<td>0.34393</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00873)</td>
<td>(0.01470)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAN LOG LIKELIHOOD</td>
<td>-1.08683</td>
<td>-1.08897</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: N=9429. Standard error in parenthesis. All test statistics are significant at 1 percent.

**TABLE 10: MODEL WITHOUT OBSERVED HETEROGENEITY**

<table>
<thead>
<tr>
<th>Model</th>
<th>Lower Trigger</th>
<th>Target</th>
<th>Upper Trigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0.74$</td>
<td>3.51</td>
<td>13.53</td>
<td>52.20</td>
</tr>
<tr>
<td>$\gamma = 1.37$</td>
<td>4.00</td>
<td>14.51</td>
<td>52.73</td>
</tr>
<tr>
<td>Eberly (1994)</td>
<td>5.92</td>
<td>12.99</td>
<td>13.16</td>
</tr>
<tr>
<td>G-L (1990)</td>
<td>5.85</td>
<td>12.99</td>
<td>29.41</td>
</tr>
</tbody>
</table>

Note: Predicted (S,s) rule is taken at the mean of each observed demographic variables, and use the fact that $E(e^\gamma) = e^{\mu x + \frac{1}{2} \sigma^2}$ if $x \sim N(\mu, \sigma^2)$. G-L stands for Grossman and Laroque.

**TABLE 11: PREDICTED (S,S) RULES**

(Percent)

<table>
<thead>
<tr>
<th>Homogeneous Model</th>
<th>Heterogeneous Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0.74$</td>
<td>$\gamma = 1.37$</td>
</tr>
<tr>
<td>6.0</td>
<td>18.2</td>
</tr>
<tr>
<td>$\gamma = 0.74$</td>
<td>$\gamma = 1.37$</td>
</tr>
<tr>
<td>4.0</td>
<td>12.8</td>
</tr>
</tbody>
</table>

**TABLE 12: PERCENTAGE OF POPULATION WHO ADJUST IN STEADY STATE**

(Annualized percent)
Figure 1: Ergodic Distribution

Appendix E: Figures

$A=1.1$, $\lambda=0.05$
Figure 3: Sample Path of Optimal \((S, s)\) Behavior
\((\alpha = 0.11 \quad \sigma = 0.59)\)
Figure 4: Ergodic Distribution and $\gamma$
Figure 5: Sample Histograms
Figure 6: Goodness of Fit Comparison
Figure 7: RDPIPC, RDEPC(M), and Recessions (Quarterly)
Figure 8: Simulated Impulse Responses
BIBLIOGRAPHY


