THREE ESSAYS IN APPLIED MICROECONOMICS

by

Carmen Eugenia Carrión Flores

A Dissertation Submitted to the Faculty of the

DEPATMENT OF ECONOMICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2007

THE UNIVERSITY OF ARIZONA

GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation

prepared by Carmen Eugenia Carrión Flores

entitled Three Essays in Applied Microeconomics

and recommend that it be accepted as fulfilling the dissertation requirement for the

Degree of Doctor of Philosophy

____Date: 4/16/07
Robert Innes

____Date: 4/16/07
Ronald L. Oaxaca

____Date: 4/16/07
Price V. Fishback

____Date: 4/16/07
John Cheslock

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

____Date: 4/16/07
Dissertation Chair: Robert Innes

____Date: 4/16/07
Dissertation Chair: Ronald L. Oaxaca

STATEMENT BY AUTHOR

SIGNED: Carmen Eugenia Carrión Flores

ACKNOWLEDGMENTS

The following dissertation, while an individual work, benefited from the insights and direction of several people. First, I would like to specially thank Dr. Robert Innes for the important support he has provided. In addition, Dr. Ronald L. Oaxaca provided invaluable advice and support at every stage of my graduate studies and of the dissertation process. Next, I wish to thank Dr. Manuela Angelucci, Dr. John Cheslock, Dr. Price Fishback, and Dr. Keisuke Hirano for giving me helpful advice and comments on earlier versions of this research. In addition, I thank my husband, Alfonso Flores-Lagunes who has provided continuos support over the years and has helped me to withstand the long process of obtaining my degree.

DEDICATION

This dissertation is dedicated to the following people since I never would have finished without their love and support.

My parents, Sergio Carrión Bogard and Carmen Flores Cano, who have been my role model for hard work, persistence and personal sacrificies, and who instilled in me the inspiration to set high goals and the confidence to achieve them.

My daughters, Carmen Eugenia and Cecilia Haydee, who had all the patience with their mother. They have grown into two wonderful ladies in spite of their mother spending so much time away from them working on this dissertation. I hope this example leads them to new and challenging horizons.

My husband, Alfonso Flores Lagunes, who has been proud and supportive of my work and who has shared the many uncertainties, challenges and sacrifices of completing this dissertation.

My parents-in-law, Alfonso Flores Sobrevilla and Haydee Lagunes Malpica, who have been very supportive and for their caring thoughts and prayers.

My grandfather, Eugenio Flores Flores to honor his name, who instilled in me the value of education.

TABLE OF CONTENTS

TABLE OF CONTENTS - *Continued*

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This dissertation applies economic theories and econometric methods to analyze the interactions between government policies and economic agents in two important and current topics: the protection of the environment and illegal migration.

Following the introduction, the second chapter studies the empirical strength of bi-directional linkages between environmental standards and performance, on the one hand, and environmental innovation, on the other. Our empirical results reveal that environmental R&D both spurs the tightening of government environmental standards and is spurred by the anticipation of such tightening, suggesting that U.S. environmental policy (at least in the context of the manufacturing industries that we study) has been responsive to innovation and effective in inducing innovation.

The third chapter studies whether a voluntary reduction pollution programs can prompt firms to develop new environmental technologies that yield future emission reduction benefits. Conversely, a VRP may induce a participating firm to divert resources from environmental research to environmental monitoring and compliance activities that yield short-term benefits in reduced emissions. We find evidence that higher rates of program participation are associated with significant reductions in the number of successful environmental patent applications four to six years after the program ended.

The fourth chapter examines the migration duration of Mexican immigrants in the U.S. using data from the Mexican Migration Project (MMP). In the past, temporary migrations were frequent, and often the rule rather than the exception in the case of Mexican immigrants. This pattern may be changing due to the tightening of the border

between Mexico and the Unites States. Moreover, this paper examines whether migration experience, demographic characteristics, economic conditions or social networks drive the time Mexican immigrants to reside illegally in the United States. The empirical analysis shows that the migration duration increases as the U.S. expected real wage increases. Tighter U.S. migration policies have an ambiguous effect on the migration duration while longer distances decrease the hazard of return to their state of origin.

In the final chapter of this dissertation, the general findings are concluded and some future avenues of research are discussed.

Chapter 1
DISSERTATION INTRODUCTION

This dissertation applies economic theories and econometric methods to analyze the interactions between government policies and economic agents in two important and current topics: the protection of the environment and illegal migration. Chapters two and three of my dissertation analyze the dynamic effects of government policies designed to provide incentives for economic agents to protect the environment. Chapter four analyzes the response of illegal migrants to some of the recent government policies intended to prevent illegal migration into the U.S.

The chapter titled "Environmental Innovation and Environmental Policy: An Empirical Test of Bi-Directional Effects" answers the following questions. First, does environmental policy spur innovation in environmental technology? Alternately, does environmental innovation lead to a tightening of environmental standards, reflecting the lower pollution abatement costs associated with better technologies? Recent empirical work focuses on the first question, finding evidence of induced innovation. In particular, higher pollution abatement expenditures (PAE) – attributable to tighter environmental policy – are estimated to increase rates of environmental patenting (Jaffe and Palmer, 1997; Brunnermeier and Cohen, 2003). However, in principle, causal effects may go in both directions: environmental policy may spur innovation, and innovation may spur tightening of environmental policy.

This observation is important for at least three reasons. First, not only does one want to understand whether and how environmental policy can yield derivative benefits in environmental innovation, but it is also of interest to understand whether regulators tighten environmental standards in response to innovation, as would be an efficient response to lowered costs of environmental compliance and would imply some sharing of the benefits of innovation with the general public. Second, even if one is only interested in "induced innovation" – how policy affects research outcomes – one needs to account for the other direction of causal effect. That is, innovation and policy are, at least in principle, jointly determined. Hence, estimates of induced innovation effects that fail to account for the joint endogeneity of innovation and policy are likely to be biased. Third, ultimately one would like to understand whether, and to what extent, tightened environmental policy can stimulate innovation and thereby yield additional long-run environmental dividends – long-run pollution reductions beyond those required by the intitial tightening of standards. To identify such benefits – the putative environmental policy multiplier – requires studying both directions of causal effect between policy and research outcomes.

To do so, we examine a panel of 127 manufacturing industries over the period 1989 – 2002 using pollutant emissions to measure policy stringency and environmental patent counts to measure environmental innovation, explicitly accounting for the joint endogeneity of environmental policy and innovation. Our empirical results reveal a negative and significant relationship between emissions and environmental patents, in both directions. Thus, environmental R&D both spurs the tightening of government environmental standards and is spurred by the anticipation of such tightening, suggesting that U.S. environmental policy (at least in the context of the manufacturing industries that we study) has been responsive to innovation and effective in inducing innovation.

The chapter titled "Do Voluntary Reduction Programs (VRPs) Spur Innovation in Environmental Technologies?" studies whether a VRP can prompt firms to develop new environmental technologies that yield future emission reduction benefits. Voluntary pollution reduction programs have become an integral part of U.S. environmental policy; there are currently over 60 partnership programs sponsored by the Environmental Protection Agency (EPA). Participants in these programs commit themselves to reduce pollutant emissions that are not addressed by environmental laws, or exceed emission standards set forth by such laws when they exist.

Economists have put forth a number of theories to explain why profit-maximizing firms self-select into costly voluntary pollution reduction programs (VPRs). Arora and Gangopadhy (1995) argue that firms want to attract a clientele of "green consumers" which are willing to pay more for goods produced in an environmentally friendly way. Voluntary pollution reductions may also deter lobbying by environmental groups for tighter environmental standards that "raise rivals' costs" (Innes and Bial, 2002); and avoid future environmental liability.

Empirical work has also documented the salutary effect of VPRs in reducing pollution (e.g., Khanna and Damon, 1999; Sam and Innes, 2006). However, to our knowledge, there has been no work to date identifying the mechanism by which VPRs may lead to emissions reductions, whether due to heightened management awareness and conscientiousness within given environmental systems and technologies or due to adoption of new environmental management systems (Anton, et al., 2004; Khanna, et al., 2005) or due to adoption of new environmental technologies.

Because pollutant reductions generally require costly reformulations of products and/or production processes, environmental over-compliance – induced by a VRP – may potentially spur environmental innovation that can reduce these costs. Conversely, a VRP may induce a participating firm to divert resources from environmental research to environmental monitoring and compliance activities that yield short-term benefits in reduced emissions. We find evidence that higher rates of 33/50 program participation are associated with significant reductions in the number of successful environmental patent applications four to six years after the program ended; these results suggest a negative relationship between the 33/50 program and longer-run environmental innovation.

The chapter titled "What Makes You Go Back Home? Determinants of the Duration of Migration of Mexican Immigrants in the US." examines the migration duration of Mexican immigrants in the U.S. using data from the Mexican Migration Project (MMP). The evolution of Mexican migration to the United States is generally understood to be the result of several forces that encourage migration. Theoretical models and recent studies on Mexican migration have suggested a dynamic pattern of cross-border migration in which the economic situation in Mexico and the United States, as well as the presence of relatives in the United States, determine the location and length of stay of Mexican migrants (Massey et al., 1987; Hanson and Spilimbergo, 1999; Lindstrom, 1996).

In the past, temporary migrations were frequent, and often the rule rather than the exception in the case of Mexican immigrants. This pattern may be changing due to the tightening of the border between Mexico and the Unites States. Moreover, this paper examines whether migration experience, demographic characteristics, economic conditions or social networks drive the time Mexican immigrants to reside illegally in the United States. The empirical analysis shows that the migration duration increases as the U.S. expected real wage increases. Tighter U.S. migration policies have an ambiguous effect on the migration duration while longer distances decrease the hazard of return to their state of origin. Finally, the period of study includes two major changes in the U.S. Immigration Law; the Immigration Reform and Control Act (IRCA) in 1986 and the Immigration Act of 1990 which are found to increase migration duration.

Chapter 5 offers a summary of this research, including a discussion of the limitations of this study. Finally, avenues for new research are outlined.

Chapter 2

DOES INDUCED INNOVATION CREATE AN ENVIRONMENTAL POLICY MULTIPLIER? AN EMPIRICAL STUDY OF BI-DIRECTIONAL LINKS BETWEEN ENVIRONMENTAL POLICY AND PATENTING[1]

2.1 Introduction

Does environmental policy spur innovation in environmental technology? Alternately, does environmental innovation lead to a tightening of environmental standards, reflecting the lower pollution abatement costs associated with better technologies? Recent empirical work focuses on the first question, finding evidence of induced innovation. In particular, higher pollution abatement expenditures (PAE) – attributable to tighter environmental policy – are estimated to increase rates of environmental patenting (Jaffe and Palmer, 1997; Brunnermeier and Cohen, 2003). However, in principle, causal effects may go in both directions: environmental policy may spur innovation, and innovation may spur tightening of environmental policy.

This observation is important for at least three reasons. First, not only does one want to understand whether and how environmental policy can yield derivative benefits in environmental innovation, but it is also of interest to understand whether regulators tighten environmental standards in response to innovation, as would be an efficient response to lowered costs of environmental compliance and would imply some sharing of the benefits of innovation with the general public. Second, even if one is only interested in "induced innovation" – how policy affects research outcomes – one needs to account

[1] Joint work with Professor Robert Innes at The University of Arizona.

for the other direction of causal effect. That is, innovation and policy are, at least in principle, jointly determined.[2] Hence, estimates of induced innovation effects that fail to account for the joint endogeneity of innovation and policy are likely to be biased. Third, ultimately one would like to understand whether, and to what extent, tightened environmental policy can stimulate innovation and thereby yield additional long-run environmental dividends – long-run pollution reductions beyond those required by the intitial tightening of standards. To identify such benefits – the putative environmental policy multiplier – requires studying both directions of causal effect between policy and research outcomes.

The purpose of this chapter is to study these bi-directional effects. Specifically, we examine 127 manufacturing industries over the fourteen-year period 1989 – 2002. Changes in environmental technologies, as measured by the number of environmental patents, can lead to changes in effective environmental standards, which in turn drive observed emissions. Emissions in turn proxy for the changes in standards that drive environmental R&D and, hence, resulting patents. In view of the joint determination of research and pollution outcomes, we estimate two simultaneous equations, using

[2] In a growing literature, economists study the links between different environmental policy instruments and innovation incentives on a theoretical level, comparing emission taxes, marketable permits, technology mandates and performance standards, with and without technology spillovers and patent protections (see Requate, 2005a). In this literature, the government typically commits to a given setting of a given regulatory instrument and allows innovation to respond accordingly. However, there is considerable anecdotal evidence that government environmental policy also *responds* to environmental innovation, often with requirements for adoption of the "best available control technology" (Jaffe, et al., 2002). Such responsive policies also provide strong incentives for environmental innovation, as they offer successful innovators a "ready market" for their products (Jaffe, et al., 2002). Innes and Bial (2002) study such responsive policies in an imperfectly competitive market setting, showing how flexible emission taxes and standards can be combined to elicit both optimal pollution levels and optimal environmental R&D (see also Requate, 2005b). With responsive policies, pollutant emissions and environmental R&D are jointly determined as successful R&D prompts policy change and attendant pollution reductions, and as anticipated policy change (and attendant tightening of pollution standards) spurs new R&D.

appropriate instruments to identify each endogenous variable.

This study contributes to a surprisingly small empirical literature on environmental innovation. This literature focuses on the effects of pollution abatement expenditures (PAE) on innovative activity. Jaffe and Palmer (1997) find evidence for the induced innovation hypothesis in U.S. industry-level panel data on total (environmental and non-environmental) R&D expenditures and patent counts. Lanjouw and Mody (1993) also find informal evidence that environmental innovation is induced by higher PAE, presenting tabular data on environmental patents and control costs from the U.S., Germany and Japan. Brunnermeir and Cohen (2003) are the first to estimate a model that links PAE to U.S. *environmental* patent counts, again finding evidence in support of the induced innovation hypothesis. [3]

Our work differs from previous studies primarily because we study a model of *bi-directional* links that explicitly accounts for the joint determination of environmental policy and environmental R&D. In doing so, we use what we consider to be a more direct measure of policy stringency, emissions as opposed to PAE. Because the two directions of causal effect are expected to be reinforcing – both negative, with higher emissions lowering research incentives, and greater research output lowering environmental standards – one expects that our accounting for joint endogeneity will dampen estimated impacts in both directions. We nevertheless find policy-induced innovation and

[3] See also related work by Popp (2002), who studies induced innovation in energy production. In addition, like us, Managi et al., (2005) are interested in bi-directional links between technology change and environmental policy stringency, in their case in the context of the offshore oil and gas industry. However, their approach is quite different than ours, examining distributed lag models of the effect of policy stringency on technology and factor productivity. We instead focus on a model of joint endogeneity in a panel of industries, building more closely upon earlier work on the induced innovation hypothesis.

innovation-induced policy effects that have the predicted negative sign, and are statistically significant.

2.2 Empirical Model

We envision an underlying structural model that determines four outcomes, our two observable variables (emissions and patents) and two unobservable variables (effective environmental standards and environmental R&D). Let us suppose that this model takes the following simple form:

$$P_{it} = a_{pit} + b_p\, RD_{it-1} + c_p\, X_{pit} + \varepsilon_{pit} \tag{2.1}$$

$$Q_{it} = a_{qit} + b_q\, S_{it} + \varepsilon_{qit} \tag{2.2}$$

$$S_{it} = a_{sit} + b_s\, P_{it} + c_s X_{sit} + d_s\, S_{it-1} + \varepsilon_{sit} \tag{2.3}$$

$$RD_{it} = a_{rit} + b_r\, \mathrm{E}_t\,(S_{it+1}) + c_r X_{rit} + d_r\, S_{it} + \varepsilon_{rit}, \tag{2.4}$$

where P_{it} is time t environmental patents in industry i, RD_{it-1} is lagged environmental R&D, Q_{it} is the volume of emissions, S_{it} is the emission standard, the vectors X_{it} represent exogenous observable variables that we describe in Section 3 below, the ε_{it}'s represent random disturbances, and E is an expectation operator. Equation (2.1) indicates that patent numbers (P_{it}) are determined by lagged industry R&D (RD_{it-1}), among other variables. Equation (2.2) indicates that emissions (Q_{it}) respond to changes in environmental standards (S_{it}). Because they are costly to firms, emission reductions beyond those required by government regulations are likely to be limited and anchored to the government's requirements; emissions are thus driven by government standards as described by equation (2.2).[4] Equation (2.3) indicates that environmental standards (S_{it}) are determined (in part) by improvements in environmental technology as measured by the number of environmental patents (P_{it}). Finally, Equation (2.4) indicates that R&D expenditures are determined (in part) by anticipated environmental standards (E_t (S_{it+1})). The impact of standards on R&D can be decomposed into two relevant effects that will be important in what follows: the impact of the anticipated *change* in environmental standards (b_r) and the impact of the initial *level* of standards (d_r+b_r). We expect the first effect to be negative and the second effect to be non-positive, as tightened (lower) emission standards promote R&D investment.

We do not have good measures of either environmental standards or environmental R&D expenditures. However, we can use relationships (2.1)-(2.4) to derive equations that indicate the relationship between environmental patents and pollutants, which we do measure. Specifically, by lagging (2.2), solving for S_{it-1} and substituting into (2.3), then substituting (2.3) into (2.2), we obtain the following structural form for emissions:

(2.5) $$Q_{it} = a^{*}_{qit} + b^{*}_{q}\, Q_{it-1} + c^{*}_{q}\, P_{it} + d^{*}_{q}\, X_{sit} + \varepsilon^{*}_{qit}$$

Intuitively, the change in environmental technology, as measured by the number of patents, drives changes in effective environmental standards, which in turn drive observed emissions. The key parameter of interest in the resulting Equation (2.5) is c_q^*, which incorporates the effects of patents on standards (b_s): $c_q^* = b_s\, b_q$, where b_q>0 from equation (2.2).

---

[4] For simplicity, other exogenous (observable) determinants of emissions are assumed to operate through standards (and the associated X_{sit} variables). At the cost of expositional simplicity, all that follows extends to the presence of other exogenous emission regressors, X_{qit}

Similarly, lagging (2.4) to substitute for RD_{it-1} in (2.1), and using (2.2) to substitute for $E_{t-1}(S_{it})$ and S_{it-1}, gives the structural form for the determination of patents:

$$(2.6) \qquad P_{it} = a^*_{pit} + b^*_p E_{t-1}(Q_{it}) + c^*_p Q_{it-1} + d^*_p X_{rit-1} + f^*_p X_{pit} + \varepsilon^*_{pit}$$

Intuitively, emissions proxy for the changes in standards that drive environmental R&D and, hence, resulting patents. The key parameters of interest in equation (2.6) are b^*_p, which incorporates the effects of anticipated policy *changes* (S_{it}-S_{it-1}) on environmental R&D (b_r), and c^*_p, which incorporates the effect of the initial *level* of standards (S_{it-1}) on R&D (d_r): $b^*_p = b_r b_p/b_q$, and $c^*_p = d_r b_p/b_q$, where b_p>0 from equation (1) and b_q>0 from equation (2.2). From our above discussion of equation (2.4), note that the *level* effect of standards on R&D is (d_r+b_r), and is proportional to the sum of the two equation (2.6) coefficients, $b^*_p + c^*_p$.

In sum, estimating equation (2.5) tests for effects of R&D on environmental policy, and estimating equation (2.6) tests for effects of environmental policy on environmental R&D. Note that emissions (from equation (2.5)) are identified by elements of X_{sit} that are not contained in the equation (2.6) set of regressors (X_{pit} and $X_{ri(t-1)}$). X_{sit} incorporates determinants of changes in "effective standards," S_{it}. As discussed below, key among such determinants are government enforcement activity that increases the stringency of environmental regulations. Likewise, patents (from equation (2.6)) are identified by elements of X_{pit} and $X_{ri(t-1)}$ that are not contained in X_{sit}. X_{pit} and $X_{ri(t-1)}$ contain variables that drive research and patent outcomes, including general trends in non-environmental research that are not relevant per se in the determination of environmental standards.

Before turning to the econometric issues relevant to the estimation of equations (2.5) and (2.6), note that equation (2.6) contains an expectation on the right hand side. The simplest (but perhaps unpalatable) way to treat this expectation is to assume that agents have perfect foresight, so that we can simply substitute the realized value Q_{it}. Then (2.5)-(2.6) give us a standard simultaneous equation framework (albeit with some complicating econometric issues that we turn to momentarily).

Now let us suppose instead that agents do not have perfect foresight. Then from (2.5)-(2.6), we have the following relationship between observable emissions and the "true regressor," $E_{t-1}(Q_{it})$:

(2.7) $$Q_{it} = E_{t-1}(Q_{it}) + u_{it}$$

where[5]

(2.8) $u_{it} = c^*_q f^*_p (X_{pit} - E_{t-1}(X_{pit})) + d^*_q (X_{sit} - E_{t-1} (X_{sit})) + \varepsilon_{uit},\quad \varepsilon_{uit} = c^*_q\ \varepsilon^*_{pit} + \varepsilon^*_{qit}$.

For our observable regressor Q_{it}, equations (2.7)-(2.8) imply two econometric problems: (1) our "true" regressor is measured with error, and (2) our observable regressor is jointly endogenous in the sense that it is correlated with the equation (2.6) error ε^*_{pit}. To obtain consistent equation (2.6) parameter estimates – addressing both of these problems – requires instruments that are uncorrelated with both the equation (2.7) "measurement error" u_{it} and the equation (2.6) disturbance ε^*_{pit} as well. Our exogenous

---

[5] Equation (2.8) follows from equation (2.5),

$$Q_{it} - E_{t-1}(Q_{it}) = c^*_q\ (P_{it} - E_{t-1}(P_{it})) + d^*_q (X_{sit} - E_{t-1} (X_{sit})) + \varepsilon^*_{qit}$$

and substitution from equation (6),

$$P_{it} - E_{t-1}(P_{it}) = f^*_p (X_{pit} - E_{t-1}(X_{pit})) + \varepsilon^*_{pit}$$

data, { X_{pit} , $X_{ri(t-1)}$, X_{sit} }, satisfies the second criterion, but unless it is all lagged, not necessarily the first. However, under the following innocuous assumption, lagged counterparts to our exogenous data satisfy both criteria:

Assumption 1. The prediction errors, $X_{pit} - E_{t-1}(X_{pit})$ and $X_{sit} - E_{t-1}(X_{sit})$, are uncorrelated with information available at time (t-1).

In what follows, we estimate equation (2.6) under both the perfect foresight premise (using contemporaneous exogenous variables and lagged instruments) and the rational expectations premise (Assumption 1, using lagged exogenous variables and instruments).

2.3 Data Description and Definition of Variables

Our sample is a balanced industry-level panel of 127 manufacturing industries (SIC codes 200-399) over the period 1989 – 2002. Because we focus on toxic emissions, we restrict attention to manufacturing industries that are the principle sources of such pollutants. Tables 2.1 and 2.2 present variable definitions and descriptive statistics for our sample.

Name of Variable	Description
SALES	Real industry sales
SALES GROWTH	Real industry sales growth measure
CONCENTRATION	Herfindahl index for each industry
CAPITAL INTENSITY	Level of new capital and equipment expenditures per-unit-sales
R&D INTENSITY	Level of research and development expenditures per-unit-sales
AGE OF CAPITAL	Net assets of an industry divided by gross assets
ENVPATENTS	Number of environmental patents, “broad” measure (Table A1)
ENVPATENTSBC	Number of environmental patents, “narrow” measure (BC, 2003)
NONENVPATENTS	Number of non-environmental patents
ENVPATENTSMA(5)	Moving average of environmental patents over the last five years
NONENVPATENTSMA(5)	Moving average of non-environmental patents over the last five years
SELFINSPECT	Number of on-site tests conducted by firms
ACTIONS	Number of enforcement actions against firms
OUTCOMP	Number of firms’ citations for out of compliance with clean air laws
EMISSIONS	Total air emissions for each industry (TRI Releases)

TABLE 2.1. Variable Definitions

Regression Sample, N= 1778 T=14		
Variables	Mean	Std. Dev
SALES	31112	103547
SALES GROWTH	-0.0348	0.2649
CONCENTRATION	0.0958	0.2197
CAPITAL INTENSITY	0.0833	0.0522
R&D INTENSITY	0.6267	0.292
AGE OF CAPITAL	0.7045	0.1429
ENVPATENTS	19.69	17.45
ENVPATENTSBC	7.507	14.761
NONENVPATENTS	21.12	22.89
ENVPATENTSMA(5)	18.47	16.19
NONENVPATENTSMA(5)	20.74	23.83
SELFINSPECT $_{t-3}$	5.17	13.43
ACTIONS $_{t-3}$	86.34	169.63
OUTCOMPLIANCE $_{t-3}$	112.81	178.34
EMISSIONS	39.473	145.073

TABLE 2.2. Summary Statistics

Using the EPA's Toxic Release Inventory (TRI), we construct industry level total toxic releases (***Emissions***) by aggregated weight by year. Facility releases reported in the TRI are assigned to the primary industry of the parent company. Following previous studies (c.f., Jaffe and Palmer, 1997; Brunnermeier and Cohen, 2003 and Popp, 2002), we use successful environmental patent applications as a proxy for environmental innovation. Using data from the U.S. Patent and Trademark Office, we construct successful patent application counts by year, by industry, environmental and non-environmental, obtained by U.S. companies.[6] Environmental patents are determined by patent classifications that relate to air or water pollution, hazardous waste prevention, disposal and control, recycling and alternative energy (***EnvPatents***). As in prior research (c.f., Jaffe and Palmer, 1997; Brunnermeier and Cohen, 2003), we determine the SIC industry to which each of these patents belongs using the primary line of business of the organization that is named first on the patent application. Table A1 in the Appendix indicates the patent utility classes that we designate as environmental in our analysis. Non-environmental patents are those in all other patent utility classes (***NonEnvPatents***). In an endeavor to include all environment-related patents in our *EnvPatents* measure, we use a broad definition of utility classes that may contain environment-related innovations. From Table 2.2, we note that our broad definition of environmental patents gives us a mean count that is almost as large as that for non-environmental patents. For robustness

[6] The literature suggests that it is preferable to count patents by date of application rather than by date of grant, because application dates better reflect the timing of discovery (uncontaminated by variability in regulatory delays). The average lag between patent applications and grants is approximately two years. All of our patent measures are for U.S. companies. U.S. companies are likely to be the most sensitive to U.S. environmental policy. Moreover, U.S. (vs. foreign) environmental innovation is more likely to be associated with an improved ability of U.S. firms to comply (at lower cost) with tightened U.S. environmental standards, and hence, to spur revisions in U.S. regulation.

purposes, we also construct a narrower measure of environmental patent counts based on the categorization of Brunnermeier and Cohen (2003); we denote this measure ***EnvPatentsBC***, and note that its sample mean is much smaller as a proportion of total patent counts (Table 2.2).

In our patent equation, we measure innovative outcomes (our dependent variable) using annual patent counts, reflecting the latest innovative responses to environmental policy. In our emission equation, however, we expect environmental standards to be revised in response to the recent history of environmental patents, not solely the last year's set of patent applications. Hence, we use a moving average of patent application counts over the preceding five years as our jointly endogenous innovation regressor; as a robustness check, we consider two alternatives as well: one and two year lagged patent counts.[7]

Our exogenous data can be broken into three categories: (1) Variables that we believe may drive both emissions and patents – that is, variables common to both X_{sit} and X_{pit}/X_{rit-1}; (2) instruments that identify emissions in the patent equation, namely, variables that are only elements of X_{sit} and not X_{pit}/X_{rit-1}; and (3) instruments that identify patents in our emission equation because they are only contained in X_{pit}/X_{rit-1} and not X_{sit}. Table 2.2 gives summary statistics for the variables that we use in our analysis. We now describe the sources and logic for our three categories of exogenous data.

---

[7] The moving average is calculated to weight more recent counts more heavily. Specifically, we use a declining balance five-year average, calculated as follows:

$$ENVPATENTSMA = \sum_{t=1}^{5} \; [(6\text{-}z)/15] \, P_{t\text{-}z},$$

where P_{t-z} is environmental patent application counts z years prior to year t.

Beginning with the first category (of common variables), we use a number of relevant financial indicators that we obtain from Standard & Poor's Compustat Services and the U.S. Department of Commerce. Deflators are obtained using producer price indexes reported in the Economic Report of the President (2004).

First, we include (deflated) industry sales volume (***Sales***) in order to account for potential effects of industry size on emissions and patents. Larger industries (ceteris paribus) are expected to produce more emissions. Expected effects on patent outcomes are less clear, as larger industries may or may not be more innovative in their environmental technologies.

Second, because market structure is a potentially important determinant of both innovative activity and environmental performance (Jaffe and Stavins, 2000; Innes and Bial, 2002), we include the four-firm Herfindahl index (***Concentration***) as an indicator of industry concentration. Expected effects of concentration on innovative activity are unclear. On one hand, more concentrated industries are more likely to be subject to the "raising rivals' costs" motives for innovative effort (Innes and Bial, 2002), with imperfectly competitive firms investing in environmental R&D in order to gain a profit-enhancing cost advantage over rival firms. On the other hand, however, firms in more concentrated industries are more likely to recognize the cost of their innovative success in prompting regulators to tighten environmental standards, thus raising their costs of environmental compliance. For example, a monopoly may avoid innovation in order to avoid higher costs of regulation. Theory also offers no clear *a priori* prediction of how concentration affects emissions. The government might regulate more concentrated industries more heavily because they are perceived to be more facile in adapting to revised standards; on the other hand, concentrated industries may be more effective at lobbying for more lax regulation.

Third, more capital intensive industries may be more polluting and have more scope for cost-reducing environmental innovation. We therefore include a measure of capital intensity (***Capital Intensity***), namely, the level of new capital and equipment expenditures divided by sales volume.

Fourth, we include each industry's total lagged level of research and development expenditures per-unit-sales (***R&D Intensity***) in order to capture effects of overall industry research activity on both environmental innovation and tightening of emission standards. Regulators may be more prone to tighten standards for more research-intensive industries that are better able to adapt (at lower cost) to regulatory changes; we therefore expect a negative coefficient on *R&D Intensity* in the emissions equation. Conversely, more research intensive industries are likely to produce environmental innovations as research byproducts (as opposed to research outcomes targeted to environmental objectives); hence, we expect a positive coefficient on *R&D Intensity* in the patent equation.[8]

Fifth, industries with older assets (ceteris paribus) may have more scope to reduce emissions and improve their environmental technology with innovation; to control for these effects, we include a measure of asset age (***Age)***, obtained by dividing total assets

[8] In principle, environmental R&D may be a component of the research intensity measure, raising the potential prospect of joint endogeneity. However, targeted environmental R&D is a very small component of overall R&D. For example, in our sample, the average annual industry-level environmental patent count calculated using the more focused measure *EnvPatentsBC* is 7.5, compared to over 40 for overall patent counts. Hence, if there is any bias, we expect it to be small and to bias against our hypothesized negative effect of environmental patents on emissions. Nevertheless, in view of this issue, we have estimated our models both with and without *R&D Intensity*, finding that our central qualitative results are robust.

of an industry by its gross assets (as in Khanna and Damon, 1999). Total assets are defined as current assets plus net property, plant and equipment and other non-current assets. Gross assets are defined as total assets plus accumulated depreciation on property, plant and equipment. *Age* is between zero and one; ratios closer to one indicate newer plant and equipment with more current assets and less depreciation.

Sixth and last, both innovation and environmental policy may be affected by the rates of growth, and hence the modernity, of the different industries. We therefore include a sales growth measure (***Salesgrowth)***.

Turning next to instruments that identify emissions (in the patent equation); we note that environmental enforcement activity is widely cited as a stimulus to pollution abatement (e.g., see Magat and Viscusi (1990), Gray and Deily (1996), Deily and Gray (2007), Decker and Pope (2006)). However, there is no evidence, in theory or empirical work, that enforcement activity affects innovative activity other than due to its effects on "effective" environmental standards and, hence, emissions.[9] We therefore use various measures of U.S. environmental enforcement activity to identify emissions. Specifically, environmental compliance and enforcement histories are obtained from the EPA's IDEA database. IDEA contains facility level data from the Aerometric Information Retrieval System (AIRS) and the Air Facility Subsystem (AFS). AFS contains compliance and

---

[9] Brunnermeier and Cohen (2003) include a measure of government environmental inspections as an explanatory variable in their patent equation. In doing so, they rightfully argue (p. 284) that "to the extent that stricter government monitoring or enforcement induces firms to comply, they might now seek less costly methods of complying." In our model, in contrast, compliance efforts (that may spur innovation) are captured by our emissions measure; that is, compliance efforts will reduce emissions, which in turn will potentially fuel environmental R&D incentives. In sum, in our paper, enforcement effects operate via emissions, even though they need not operate via PAE, the policy proxy in Brunnermeier and Cohen's (2003) analysis.

enforcement data on stationary sources of air pollution. Regulated sources range from large industrial facilities to relatively small operations. We use counts of enforcement actions (***Actions***), numbers of facilities out of compliance with clean air laws (***Outcomp***), and the number of reported self-inspections (***Selfinspect***) as indicators of environmental enforcement stringency. Because enforcement effects on emission performance occur with a substantial delay, we lag all of our instruments by three years.[10] For robustness purposes, we consider a variety of different instrument combinations; we report results using two combinations but have obtained similar results using other instrument menus.

To identify environmental patent counts in our emission equation, we use corresponding (moving average or lagged) non-environmental patent counts. Intuitively, trends in overall innovative output are reflected in a high correlation between these two patent measures; for example, environmental and non-environmental patents by U.S. companies have a correlation coefficient equal to .75 in our sample. On the other side of the coin, is there any reason to expect non-environmental patents to be relevant to the determination of emissions (other than via effects on environmental patenting)? In principle, there may be two reasons (that we can think of), and we control for both. First, perhaps there are effects of overall research proficiency on the economic adaptability of different industries to regulatory changes, which in turn influence regulatory standard setting; we control for such effects by including lagged *R&D Intensity* as a regressor. Second, perhaps non-environmental innovation increases overall industry productivity,

---

[10] Lagging three years has the added advantage of avoiding any potential for endogeneity between emissions and enforcement. We considered other enforcement lags and found that three-year lags in our three enforcement variables performed the best as determinants of emissions.

and hence output, thus raising emissions; we control for such effects by including an industry output measure (***Sales***) as a regressor.[11]

As always, two key criteria underpin our instrument choices. First, the instruments should be highly correlated with the jointly endogenous variable that they identify. In linear simultaneous systems, a common statistical test for this property is obtained from first stage regressions of the endogenous variables on all exogenous data. In our emissions equation, however, we have a lagged dependent variable (and evidence of serial correlation when treating the lag as exogenous); hence, we perform both a standard first-stage regression (on purely exogenous data) and a dynamic panel analog to the "first-stage" regression (following Arellano and Bover, 1995, and Blundell and Bond, 1998, as discussed in detail in the next section). Table 3 reports estimates for the pure and pseudo (dynamic) first stage models for our emission equation. In all cases, note that our identifying instruments, *Selfinspect*, *Outcomp* and *Actions*, are jointly significant. We expect (from prior work and intuitive logic) that lagged enforcement scrutiny, as measured by enforcement actions and compliance status, will spur reductions in emissions. In contrast, we expect that self-inspections may substitute for government scrutiny and, hence, favor laxity in emissions performance. The "first stage" estimations in Table 2.3 are consistent with these expectations.

[11] A potential concern with use of non-environmental patents as an identifying instrument is that we may improperly classified some "environmental" patents as "non-environmental." For this reason, we make our definition of "environmental" utility classes broad, incorporating all classes that have potential environmental relevance (see Appendix). As a result, our EnvPatents variable has a mean almost three times that of the more narrow measure of Brunnermeier and Cohen (2003) (see Table 2.2).

Dependent Variable	Emissions							
Variable Instrumented	None				Emissions t-1			
	Model 1: Fixed Effects		Model 2: Fixed Effects		Model 3: Dynamic Model		Model 4: Dynamic Model	
Variables	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	z	Coefficient (Robust SE)	z
SELFINSPECT t-3	17.987*** (4.7929)	3.75	17.143*** (4.9841)	3.44	8.633*** (2.789)	3.10	15.551*** (3.512)	4.43
OUTCOMP t-3	-234.50*** (60.296)	-3.89	-196.90*** (85.6117)	-2.30	-43.892*** (19.491)	-2.25	-61.037** (27.2356)	-2.24
ACTIONS t-3	*	*	-56.344* (30.0557)	-1.87	*	*	-36.576* (18.856)	-1.94
R&D INTENSITY	-0.0358 (0.1119)	-0.32	-0.0762 (0.3268)	-0.23	-0.0025 (0.0057)	-0.44	-0.0017 (0.0033)	-0.52
CAPITAL INTENSITY	127.889 (79.240)	1.61	126.865 (79.277)	1.60	85.892** (53.412)	1.61	63.241 (61.3990)	1.03
CONCENTRATION	154.136 (1760.97)	0.09	102.949 (176.335)	0.06	56.561* (30.908)	-1.83	51.913** (26.486)	1.96
AGE OF CAPITAL	51.086 (38.518)	1.33	50.026 (38.5662)	1.30	-44.427 (83.825)	-0.53	-38.273 (60.750)	-0.63
SALES	0.0722 (0.1602)	0.45	0.0802 (0.1607)	0.50	0.0190** (0.0096)	1.98	0.0242* (0.013)	1.86
SALES GROWTH	-3.6020 (11.6060)	-0.31	-3.64 (11.609)	-0.31	-3.299 (2.750)	-1.20	-3.406 (2.805)	-1.21
NONENVPAT	-0.8378* (0.4429)	-1.89	-0.8476* (0.4429)	-1.91	-0.5191*** (0.2579)	-2.01	-0.6019** (0.2487)	-2.42
EMISSIONS t-1	*	*	*	*	0.3473 (0.2381)	1.46	0.3804* (0.1957)	1.94
CONSTANT	7.1342 (33.6860)	0.21	8.732 (33.794)	0.26	8.643 (7.463)	1.16	9.732 (8.145)	1.19
R-sq (with instruments)	0.3476		0.3479		*		*	
R-sq (without instruments)	0.1192		0.1219		*		*	
					Statistic	p-value	Statistic	p-value
Hansen Test	*		*		7.07	0.422	7.23	0.3
AR(1)	*		*		-1.34	0.179	-1.34	0.179
AR(2)	*		*		0.88	0.378	0.88	0.381

TABLE 2.3. “First Stage” Estimation Results for patent equation

Similarly, Table 2.4 provides statistical evidence of the “first stage” relationship between environmental patent counts and non-environmental patent counts. Here, we present both linear and Poisson fixed effects estimations of “first stage” patent equations. Again, we find that our identifying instrument (non-environmental patents) is a significant predictor of environmental patent measures, with the predicted positive sign.

	Model 1: Fixed Effects		Model 2: Fixed Effects		Poisson FE		Poisson FE	
Dependent Variable	**USENVPATMA5**				**USENVPAT t-1**		**USENVPAT t-2**	
Variables	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	z	Coefficient (Robust SE)	z
USNONPATMA5	0.0614*** (0.0056)	10.89	0.0768*** (0.0071)	10.68	*	*	*	*
NONENVPAT t-1	*	*	*	*	0.0317*** (0.0058)	5.47	*	*
NONENVPAT t-2	*	*	*	*	*	*	0.0273*** (0.0073)	3.74
R&D INTENSITY	0.0005** (0.0002)	2.45	0.0004** (0.0002)	1.96	0.0001** (0.00004)	2.5	0.0017*** (0.0004)	4.25
CAPITAL INTENSITY	18.29 (47.13)	0.39	28.88 (58.52)	0.49	26.6038 (58.3284)	0.46	23.694 (49.573)	0.48
CONCENTRATION	-39.43 (80.55)	-0.49	-41.61 (80.52)	-0.51	-37.46 (81.538)	-0.46	-32.428 (81.05)	-0.40
AGE	8.89 (20.02)	0.44	18.29 (28.74)	0.64	14.4584 (28.3294)	0.51	13.648 (28.8827)	0.47
SALES	0.0002* (0.0001)	1.91	0.0002* (0.0001)	1.67	0.0003 (0.0002)	1.5	0.0002* (0.00014)	1.63
SALES GROWTH	-1.7839 (6.1945)	0.29	-1.7539 (6.1972)	-0.28	-1.4527 (7.5284)	-0.19	-1.2547 (7.7365)	-0.16
SELFINSPECT t-3	-0.37 (0.9048)	-0.4	-0.1734 (0.3672)	-0.47	-0.4921 (0.5942)	-0.83	-0.4393 (0.6011)	-0.73
OUTCOMP t-3	0.0509 (0.386)	0.13	0.0143 (0.0632)	0.23	0.1185 (0.1055)	1.12	0.1104 (0.1062)	1.04
ACTIONS t-3	*	*	0.0394 (0.0671)	0.59	*	*	*	*
CONSTANT	45.4319** (18.1277)	2.51	44.69023** (18.4239)	2.43	*	*	*	*
R-sq (with instruments)	0.1786		0.1798		*		*	
R-sq (without instruments)	0.1276		0.1127		*		*	
Log-Likelihood	*		*		-22615.561		-23788.412	

TABLE 2.4 “First Stage” Estimation results for emission equation

Second, the instruments for emissions (patents) should be uncorrelated with the errors in the patent (emission) equation. Beyond our intuitive arguments that there is no correlation, the best we can do to test for this property is to examine the validity of our over-identifying restrictions. Corresponding (Hansen / Sargan) test statistics are constructed for each estimated equation and reported in the tabular results of Section 5 below. Note that, in all cases, we do not reject our maintained (null) hypothesis of no correlation (with p-values above twenty percent in almost all cases).

2.4 Econometric Methods

We have two simultaneous equations which we estimate equation-by-equation.[12] In doing so, a variety of econometric issues arise. First, we have a panel data structure and, hence, need to account for individual effects. Second, we have endogenous regressors. Third, our emission equation has a dynamic structure. And fourth, our observed patent measure takes a count form for which we must account in our estimation strategy. In what follows, we describe how we address these issues in each of the two equations.

2.4.1 Emission Equation Estimation Technique

Our econometric analysis of the emission equation is based on equation (5), with industry fixed effects.[13] The disturbance term, ε^*_{qit} is assumed to be independently

[12] In principle, one can gain some efficiency if the two equations are estimated as a system. However, we prefer to estimate equation by equation for simplicity (given that we have a distinct set of estimation issues for each equation) and in order to avoid any potential bias due to any cross-equation misspecification.

[13] Formally, we assume that $a^*_{qit} = \lambda_{qt}+\mu_{qi}$. Because the time dummies are found to be jointly insignificant, they are dropped from the estimation for the sake of efficiency.

distributed across industries with zero conditional mean. However, no restrictions are placed on heteroskesdasticity across industries and time.[14]

Because we have a dynamic linear panel model, standard estimators that ignore the lagged dependent variable, or fail to account for its potential endogeneity, are biased and inconsistent (Baltagi, 1995). Arellano and Bond (1991) are the first (to our knowledge) to propose a Generalized Method of Moments (GMM) estimator for a dynamic panel data model with endogenous regressors that is consistent (in the number of cross-section units) for a fixed time horizon. Arellano and Bover (1995) and Blundell and Bond (1998) subsequently recommend more efficient estimators. In particular, Blundell and Bond (1998) develop a system GMM estimator with a two-step finite sample correction (see also Windmeijer, 2000). We use the system GMM variant mainly because the two-step estimator uses a weighting matrix which is (asymptotically) efficient and heteroskedasticity consistent.[15]

Because most estimates of emission equations in the literature are based on static models, we also want to compare our estimates to those obtained with traditional static methods (i.e., a model without lagged emissions on the right hand side). Therefore, we also present a non-dynamic (fixed effects) IV estimation.

2.4.2 Patent Equation Estimation Technique

So far, in deriving our patent equation (2.6), we have implicitly assumed a linear

[14]In estimating (5), we considered a variety of alternative lag structures for both Q and the exogenous data. In all cases, we could not rejct the null hypothesis that additional lags of Q and X are equal to zero; p-values for these hypotheses range from 0.2384 to 0.6145

[15] This matrix is calculated using the estimated residuals from the one-step estimation; see Arellano and Bond (1991).

process that generates a continuous variable. However, measured patent outcomes take a count form, with no negative values, a substantial number of zeroes (roughly one third in our sample), and integer positive values that range from one to 153 (with half of the positive values less than 40). Conceptually, we interpret patent outcomes as the observable consequence of our continuous (and unobservable) index of technology change P_{it} (of equation (2.6)). Specifically, let us suppose that patent counts P^*_{it} are distributed Poisson with

$$E(P^*_{it} / \varepsilon^*_{pit}) = exp(P_{it}), \qquad (2.9)$$

where P_{it} is determined by equation (2.6) with industry fixed effects.[16] This gives us the multiplicative error Poisson panel model, with endogenous regressors, of Blundell, Griffith and Windmeijer (2002) (see also Windmeijer (2002) and Windmeijer and Santos Silva (1997)). This is the model we use to estimate our patent equation.[17]

2.5 Empirical Findings

Before turning to our two equations, we note that a key issue motivating our work is the prospective joint endogeneity of emissions and patent outcomes. Given endogeneity tests available to us, we are able to provide some preliminary evidence that we indeed have simultaneity in our sample. In particular, for our IV fixed effects emissions

[16] As in the emission equation, we allow for both time and industry fixed effects. However, the time dummies are again jointly insignificant; hence, for efficiency, we estimate with industry fixed effects only.
[17] Because we have a mixture Poisson with multiplicative error, our estimation allows for over-dispersion (see Cameron and Trivedi, 1998, p. 98) and thus avoids the main criticism of a standard fixed effects Poisson. To our knowledge, there is no known Negative Binomial counterpart to the Poisson estimator of Blundell, et al. (2002) and Windmeijer and Santos-Silva (1997) that accounts for our case of an endogenous regressor with a nonlinear (dynamic) generating process.

equation, we can test for the endogeneity of patents (***Usenvappma5***) with a standard Hausman statistic; the resulting (Chi-square (1)) statistic is 15.91 with a p-value of 0.0002, clearly rejecting the null of exogeneity in the patent variable. In the patent equation, we can also construct a Hausman statistic provided we restrict the model to have only contemporaneous emission effects (see Grogger, 1990; and Windmeijer and Santos Silva, 1997); doing so for one of our main patent models (our Rational Expectations Model 2 of Table 2.7 below) yields a test statistic equal to 5.49 with a p-value of 0.001.[18] Again, we clearly reject the null of exogeneity in the emission variable.[19] Both statistics indicate a need to account for endogeneity of emissions and patents in both equations.

2.5.1. Emission Equation Implications

Table 2.5 and 2.6 present estimation results for the dynamic panel model of the emission equation (2.5). Four dynamic panel estimations are presented, with two alternate sets of enforcement measures, and three alternative measures of lagged environmental patent counts: Lagged five year moving average of environmental patents (which we view as our best measure), one-year lagged counts, and two-year lagged counts. Note that test statistics for serial correlation (m_1 and m_2) and overidentifying restrictions (Hansen) do not indicate misspecification in any of the models.[20] The coefficient for the lagged

---

[18] Corresponding Hausman statistics / p-values for our other models are 5.23 / .001 (Model 1, Table 2.6A), 4.92 / .002 (Model 2, Table 2.6A), and 5.71 / .001 (Model 1, Table 2.6B).

[19] These tests are clearly only illustrative as they fail to account for the dynamic (lagged) effects of emissions in either equation.

[20] The test statistics m1 and m2 test for the presence of serial correlation in the first differenced residuals of first and second order, respectively, asymptotically distributed as a N(0,1) under the null hypothesis of no

dependent variable is 0.7174 using Model 3, and is statistically significant.[21] Performing the unit root test developed by Levin, et al. (2002), we reject the null hypothesis that the emissions series contains a unit root, thus indicating that the series is stationary.[22]

---

serial correlation (see Arellano and Bond, 1991). As expected, there is significant negative first order autocorrelation, but no significant second order autocorrelation, a crucial property for the validity of our instruments. Moreover, the Hansen (1982) test statistic for overidentifying restrictions is □2 –distributed with degrees of freedom equal to the number of instruments minus the number of estimated parameters. This misspecification test does not indicate correlation between the instruments and the error term. We report the Hansen test statistic rather than the Sargan (1958) test statistic because it is robust to heteroskedasticity and autocorrelation. For a more detailed discussion, see Hansen (1982), Hansen and Singleton (1982), and Newey and West (1987).

[21] This estimated coefficient lies in the interval between the within group and OLS estimates (of 0.5665 and 0.893, respectively), as expected.

[22] The Levin statistic for Model 3 is -0.5452 with a t-value of -33.17.

	Model 1: IV Fixed Effects		Model 2: Dynamic Model		Model 3: Dynamic Model	
Dependent Variable	EMISSIONS					
Variable Instrumented	EMISSIONS t-1 and ENVPATENTSMA(5)					
Variables	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t
SELFINSPECT t-3	18.29*** (5.0806)	3.60	13.66*** (2.7)	5.05	16.4*** (3.53)	4.64
OUTCOMP t-3	-22.21*** (6.7920)	-3.27	-40.44*** (20.01)	-2.02	-52.47*** (25.78)	-2.03
ACTIONS t-3	*	*	*	*	-24.7** (12.7)	-1.89
ENVPATENTSMA(5)	-0.4609*** (0.1344)	-3.43	-0.1077*** (0.048)	-2.20	-0.1034*** (0.047)	-2.20
ENVPATENTS t-1	*	*	*	*	*	*
ENVPATENTS t-2	*	*	*	*	*	*
EMISSIONS t-1	*	*	0.7166*** (0.0588)	12.18	0.7174*** (0.0582)	12.32
R&D INTENSITY	-0.351 (0.3529)	-0.99	-0.1119*** (0.0372)	-3.00	-0.1204*** (0.0379)	-3.18
CAPITAL INTENSITY	11.783 (8.4164)	1.40	93.8 (65.09)	1.44	87.81 (63.21)	1.37
CONCENTRATION	-12.927 (184.671)	-0.07	-10.19*** (3.98)	-2.56	-8.39*** (3.89)	-2.15
AGE OF CAPITAL	6.1543 (4.1304)	1.49	-6.38*** (2.76)	-2.31	-5.3*** (2.61)	-2.03
SALES	0.2138 (0.2076)	1.03	0.0411*** (0.0149)	2.75	0.0409*** (0.0148)	2.76
SALESGR	-7.645 (12.3306)	-0.62	-12.59*** (4.43)	-2.84	13.52*** (5.38)	-2.55
CONSTANT	37.814 (36.713)	1.03	47.29 (62.39)	0.76	53.07 (57.46)	0.92
	Statistic	p-value	Statistic	p-value	Statistic	p-value
Hansen Test	*	*	41.9	0.431	45.58	0.365
AR(1)	*	*	-1.74	0.082	-1.74	0.082
AR(2)	*	*	0.44	0.658	0.43	0.665

TABLE 2.5 Emission Equation Estimation Results

	Model 4: Dynamic Model		Model 5: Dynamic Model	
Dependent Variable	**EMISSIONS**			
Variable Instrumented	**EMISSIONS t-1 and ENVPATENTS t-1**		**EMISSIONS t-1 and ENVPATENTS t-2**	
Variables	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t
SELFINSPECT t-3	19.05*** (4.34)	4.38	10.29*** (3.65)	2.81
OUTCOMP t-3	-52.50*** (25.8)	-2.03	-52.63*** (25.75)	-2.04
ACTIONS t-3	*	*	*	*
ENVPATENTSMA(5)	*	*	*	*
ENVPATENTS t-1	-0.1383*** (0.0576)	-2.4	*	*
ENVPATENTS t-2	*	*	-0.0676*** (0.0295)	-2.29
EMISSIONS t-1	0.7175*** (0.0582)	12.32	0.7173*** (0.0583)	12.3
R&D INTENSITY	-0.1214*** (0.0372)	-3.26	-0.1218*** (0.0377)	-3.23
CAPITAL INTENSITY	84.35 (63.58)	1.32	85.67 (63.36)	1.35
CONCENTRATION	-8.58*** (3.89)	-2.2	-8.67*** (3.36)	2.58
AGE OF CAPITAL	-5.78** (2.97)	-1.94	-6.28*** (2.97)	-2.11
SALES	0.0411*** (0.0142)	2.89	0.041*** (0.0145)	2.82
SALESGR	-13.26*** (2.45)	-5.41	-14.26*** (2.46)	-5.79
CONSTANT	56.77 (57.44)	0.98	53.5 (57.53)	0.92
	Statistic	p-value	Statistic	p-value
Hansen Test	45.44	0.412	43.81	0.437
AR(1)	-1.74	0.082	-1.74	0.082
AR(2)	0.44	0.663	0.43	0.664

TABLE 2.6 Emission Equation Estimation Results

Qualitative implications of Tables 2.5 and 2.6 can be summarized as follows.

1) *Technological innovation spurs a tightening of emission standards*. In all specifications – and with all three alternative measures of technological progress / patent counts – we find negative and significant effects of environmental innovation on emissions. We interpret such costly intra-industry emission reductions to imply a corresponding tightening of toxic emission standards, as firms will surely not engage in costly emission abatement that is not otherwise required.[23] Assessing the quantitative importance of these effects is not particularly easy. For example, Model 3 implies that the estimated long-run effect of one patent (approximately 5.4 percent of the sample mean) is to reduce associated industry emissions by 0.2 percent (of sample mean).[24] Although the marginal effects of patents on emissions are small, the total effects of patents, when taken cumulatively, are indeed significant. Our Model 3 estimates imply, for instance, that one year of innovative success (evaluated at the sample mean of the moving average of environmental patents) spurs a 3.8 percent long-run reduction in emissions.

2) *Emission standards tend to be tighter for industries that are more concentrated, have newer assets, and are growing more rapidly*, with significant negative coefficients on our measures of concentration, asset age, and sales growth. All of these effects are consistent with the hypothesis that regulators impose tighter standards in industries that

[23] In principle, if cross-plant emissions trading were possible, there could be an alternative interpretation of our results: Improved industry-level environmental technology (as measured by a higher patent count) may spur emission permit sales from the innovating industry to other industries. However, for the hazardous pollutants that are reported in the TRI, U.S. regulation does not allow cross-plant trading of emission rights (see, for example, U.S. Code, Title 42, Section 7412). Hence, emission reductions are net (i.e., not offset in other industries) and thus represent tightening of industry-level emission standards.

[24] This percentage is obtained by converting Model 3 into difference form (subtracting lagged emissions from both sides) and solving for the long-run marginal effect of a patent on the change in emissions.

are deemed to be more facile (i.e., better able at lower cost) to adapt to stronger regulation.

2.5.2. Patent Equation Implications

Tables 2.7 and 2.8 present estimation results for our patent equation. Table 2.7 presents results under a perfect foresight premise that next period emission standards are foreseen by industry participants; hence, regressors can be contemporaneous (see Section 2.2 above). Table 2.8 presents results under the alternative rational expectations (Assumption 1) premise, requiring that exogenous variables be lagged. We use a two-year lag, assuming that R&D investments (in equation (4)) are driven by a two-year-ahead policy forecast. From both Tables, note that test statistics for serial correlation (m_1 and m_2) and over-identifying restrictions (Sargan) do not indicate misspecification.

In each Table, we present four models, two each using our broad environmental patent measure *EnvPatents* and our more focused measure *EnvPatentsBC*, respectively. In each case, we report models with two alternative instrument sets to identify emissions. Moreover, note that in both cases we measure non-environmental patent counts using our broad measure of environment-related patents; we do this to ensure that the non-environmental patent regressor is uncontaminated by any potential environment-related research outcomes.

	Model 1		Model 2		Model 3		Model 4	
Dependent Variable	**ENVPATENTS**				**ENVPATENTSBC**			
Variable Instrumented	**Emissions and Emissions t-2**							
Variables	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t
EMISSIONS	-0.00943*** (0.001608)	-5.86	-0.00968*** (0.001459)	-6.63	-.01024* (0.0057714)	-1.77	-.01122** (0.005062)	-2.22
EMISSIONS t-1	0.004242*** (0.00022)	19.09	0.003123*** (0.000256)	12.18	0.003994** (0.00201)	1.98	0.003918** (0.001656)	2.37
R&D INTENSITY t-1	8.8660*** (1.3450)	6.58	8.245 (5.388)	1.53	6.3870* (3.8546)	1.66	6.3821* (3.7216)	1.71
CAPITAL INTENSITY	8.341*** (0.715)	11.66	9.115*** (2.476)	3.68	8.860*** (2.902)	3.05	9.0343*** (2.9008)	3.11
CONCENTRATION	-13.833*** (1.32)	-10.45	-13.298** (5.2)	-2.55	-14.184*** (4.43)	-3.20	-14.0392*** (4.3647)	-3.20
AGE OF CAPITAL	-0.497** (0.246)	-2.02	-0.5401 (4.06)	-0.13	-2.272 (4.536)	-0.50	-2.2667 (4.5698)	-0.50
REAL SALES	-19.6200*** (3.9900)	-4.91	-18.498 (17.11)	-1.08	-13.8492 (8.9194)	-1.55	-13.1132 (8.5812)	-1.53
SALES GROWTH	0.1314*** (0.0428)	3.06	0.0678 (0.0618)	1.09	0.2437 (0.6418)	0.37	0.2138 (0.6485)	0.32
NONENVPAT	0.0051*** (0.0004)	13.09	0.0051* (0.0027)	1.88	0.0051** (0.0023)	2.18	0.0050** (0.0023)	2.20
Instruments								
SELFINSPECT t-3	YES		YES		YES		YES	
OUTCOMP t-3	YES		YES		YES		YES	
ACTIONS t-3	NO		YES		NO		YES	
					Statistic	p-value	Statistic	p-value
Hansen Test	24.37	0.2262	26.35	0.1545	23.8025	0.2511	23.7915	0.2043
AR(1)	-1.56	0.118	0.1088	0.1180	-1.6214	0.1049	-1.5546	0.1200
AR(2)	-0.7659	0.4438	0.5863	0.4438	-0.6284	0.5297	-0.8118	0.4169

TABLE 2.7. Patent Estimation Results under Perfect Foresight

	Model 1		Model 2		Model 3		Model 4	
Dependent Variable	**ENVPATENTS**				**ENVPATENTSBC**			
Variable Instrumented	**Emissions and Emissions t-2**							
Exogenous Variables	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t	Coefficient (Robust SE)	t
EMISSIONS	-0.00834*** (0.001114)	-7.48	-.01243*** (0.001608)	-7.72	-.010041 (0.006747)	-1.49	-.010844* (0.006056)	-1.79
EMISSIONS t-2	0.002130*** (0.00027)	7.87	0.004242*** (0.00022)	19.02	0.003804*** (0.001298)	2.93	0.003663*** (0.001382)	2.65
R&D INTENSITY t-2	6.6620*** (0.5108)	13.04	8.866*** (1.3458)	6.58	6.6854** (3.1837)	2.10	6.2484 (4.0028)	1.56
CAPITAL INT. t-2	4.122*** (0.732)	5.63	8.341*** (0.7149)	11.66	8.806*** (2.342)	3.76	8.8979*** (2.4823)	3.58
CONCENT t-2	-12.790*** (0.74)	-17.34	-13.83*** (1.323)	-10.45	-13.656** (6.923)	-1.97	-13.2261* (7.086)	-1.87
AGE OF CAPITAL t-2	-0.504 (0.569)	-0.88	-0.497** (0.246)	-2.02	-1.902** (0.7543)	-2.52	-1.746*** (0.673)	-2.60
SALES t-2	-11.5900*** (1.3620)	-8.50	-19.62*** (3.993)	-4.91	13.0183* (7.2394)	1.80	11.6192 (7.301)	1.59
SALES GROWTH t-2	0.0139 (0.0852)	0.16	0.1314*** (0.0428)	3.06	0.2332 (0.1629)	1.43	0.2056*** (0.0642)	3.20
NONENVPATENTS t-2	0.0049** (0.0020)	2.45	0.0052** (0.0026)	1.99	0.0047 (0.0029)	1.62	0.0045* (0.0026)	1.73
Instruments used								
SELFINSPECT t-3	YES		YES		YES		YES	
OUTCOMPT t-3	YES		YES		YES		YES	
ACTIONS t-3	NO		YES		NO		YES	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Hansen Test	25.35	0.1881	24.37	0.2262	24.28	0.2215	23.8956	0.2319
AR(1)	-1.1377	0.2552	0.118	0.2552	-1.3294	0.2561	-1.3554	0.2554
AR(2)	-0.8027	0.4221	0.4438	0.4221	-0.8513	0.4829	-0.8437	0.4365

TABLE 2.8. Patent Estimation Results under Rational Expectations

Key qualitative implications of our results can be summarized as follows.

1) *Environmental innovation is spurred by the anticipated tightening of emission standards*. As noted in Section 2.2, we are interested in two effects of emissions standards on R&D. The first is the impact of the anticipated *change* in standards and is measured by the coefficient on contemporaneous emissions. The second is the effect of the initial *level* of standards and is measure by the sum of the coefficients on contemporaneous and lagged emissions.

Turning to the first effect, we see that, in all models, the estimated coefficient on emissions is negative and significant; hence, anticipated reductions in industry-level emissions standards lead to increases in successful patent applications. In quantitative terms, these estimated effects are of roughly similar magnitudes across the different models and different environmental patent measures. Similar estimated coefficients across the two environmental patent measures imply that proportional impacts of changes in environmental standards are similar for the two measures. Hence, environmental patents included in our "broad" measure (*EnvPatents*), but not included in our "narrow" (Brunnermeier and Cohen, 2003) measure (*EnvPatentsBC*), have similar sensitivity to environmental policy as do those in the "narrow" category. To assess the magnitude (and economic importance) of these effects, consider our Rational Expectations (Table 2.6) Model 2; in this Model, a one percent (of sample mean) reduction in anticipated emissions is estimated to increase successful environmental patent applications by roughly one-half of one percent (0.49).

With regard to the second (level) effects of standards, we again see estimated effects of roughly similar magnitudes across the different models, again implying similar proportional impacts for our two environmental patent measures. In all cases, as expected, initial emission standards have significant negative effects on R&D / patent outcomes, implying that tighter initial standards spur environmental R&D. To assess the magnitude of the estimated effects, let us again use our Rational Expectations Model 2 (of Table 2.8) to illustrate; in this Model, a one percent reduction in the initial level of emissions (based on the sample mean of emissions) is estimated to increase subsequent environmental patenting by roughly one-third of one percent (0.32, based on the sum of the two emissions coefficients).

The magnitudes of our "induced innovation" effects are large by comparison to earlier work. Brunnermeier and Cohen (2003), for example, find that a one percent increase in pollution abatement costs spurs an increase in successful environmental patent applications of approximately four-one-hundredths of one percent. The larger impacts that we find are likely due to our different (emission-based) measure of policy stringency. However, our estimated effects are nonetheless rather small in the following sense.

2) *The "multiplier effect" of induced innovation on long-run emissions – what we have termed the "environmental policy multiplier" – is proportionately small.* Consider the impact of an exogenous one percent (of sample mean) permanent tightening in emission standards. Simulating resulting changes in emissions and patents over time using Model 3 of Table 2.5 and Model 2 of Table 2.8, we estimate an additional long-run emission reduction of 2.74 percent and a long-run increase in annual environmental patenting of 1.2 percent (as shares of sample average emissions and patents,

respectively). A key question here is: How much of the additional emission reduction is attributable to the additional patenting? The answer (obtained by comparing to simulated outcomes with no induced innovation) is 0.08182, which is 7.55 percent of the additional (2.74 percent) emission reduction and 20.7 percent of the initial (one percent) emission reduction.[25] While this impact is not inconsequential, it is also not particularly large.

In sum, we find that tightened emission standards spur environmental innovations that in turn fuel greater emission reductions. However, the proportionate contribution of *induced* innovation to long-run emission reduction appears to be modest. On the other hand, the contribution of *overall* innovation to long-run emission reductions is estimated to be substantial (Table 2.5). It would thus appear that environmental innovation, stimulated in part by environmental policy but predominantly by overall technological advancement, is a very important driver of progress in ultimate pollution reduction.

3) *Environmental innovation tends to be greater in more research intensive, more capital intensive, more rapidly growing, smaller, and less concentrated industries*. Intuitively, more capital intensive industries with older assets may have more scope and incentive for emission-reducing innovation; notably, this result is consistent with prior work that finds innovation incentives to rise with capital intensity and pollution abatement expenditures that are higher when assets are older. Larger and more concentrated industries may better internalize prospective costs of innovation in leading regulators to tighten environmental standards, costs that can deter innovation. Potentially, smaller and less concentrated industries may also be more innovative by nature, and be

---

[25] The remainder is attributable to the dynamic multiplier.

able to distinguish themselves in "green markets" as environmentally proactive corporate citizens (Arora and Cason, 1996). More rapidly growing and more research intensive industries, as expected, are more active in environmental patenting.

2.6 Conclusions

In this chapter, we presented empirical evidence of bi-directional linkages between environmental standards and environmental performance, on the one hand, and environmental innovation, on the other. Pollutant emissions and environmental R&D are jointly determined as successful R&D prompts policy change and attendant pollution reductions, and as anticipated policy change (and attendant tightening of pollution standards) spurs new R&D. Specifically, we examined 127 manufacturing industries over the fourteen-year period 1989 – 2002, accounting for the joint determination of research and pollution outcomes.

Our empirical results reveal a negative and significant relationship between emissions and environmental patents, in both directions. Thus, environmental R&D both spurs the tightening of government environmental standards and is spurred by the anticipation of such tightening. Empirical results also suggest that a linear feedback model is appropriate in order to capture the dynamic nature of the links between environmental policy and environmental innovation.

These results suggest that there is a salutary process by which the promise of tightened standards stimulates environmental research, and environmental research, by lowering costs of abatement, stimulates tighter standards. However, the ultimate benefits of tightened pollution standards, due to the resulting stimulus to environmental innovation, appear to be modest. While environmental innovation is found to be a very important driver of long-run pollution reduction, environmental policy plays a role in stimulating environmental research that is statistically significant and not inconsequential, but proportionately not very large.

Moroever, our results say nothing about the efficiency of environmental policy in stimulating research. Indeed, these results are consistent with (but do not imply) a regulator who chooses standards that are ex-post efficient – that is, efficient for any given state of technology – but not chosen with ex-ante commitments that account for impacts on research incentives (see Requate, 2005b; Innes and Bial, 2002; Innes, 2006). Hence, there is no evidence per se that regulators set tighter standards – vis-à-vis those that are ex-post efficient – in order to spur more innovation, as one might interpret Michael Porter's (1990) famous conjecture to imply.

This observation, as well as the aggregations we make in this study, suggests natural avenues for further inquiry. For example, how do different forms of regulation – tighter standards vs. voluntary pollution reduction programs vs. updated technological regulations – affect innovative effort? And how do different types of innovative effort (more exploratory vs. more derivative) influence and get influenced by environmental standards and regulation? Finally, is there any sense in which regulatory strategy is optimal in inducing and responding to environmental innovation? All of these issues, we believe, merit further study.

Chapter 3

DO VOLUNTARY POLLUTION REDUCTION PROGRAMS (VPRS) SPUR INNOVATION IN ENVIRONMENTAL TECHNOLOGY?[26]

3.1 Introduction

Voluntary pollution reduction programs have become an integral part of U.S. environmental policy; there are currently over 60 partnership programs sponsored by the Environmental Protection Agency (EPA). Participants in these programs commit themselves to reduce pollutant emissions that are not addressed by environmental laws, or exceed emission standards set forth by such laws when they exist. Current partnership programs include ``AgStar" which promotes the use of biogas recovery systems to curb methane emissions at confined animal feedlot operations and ``EnergyStar" which seeks to reduce carbon dioxide emissions.

Economists have put forth a number of theories to explain why profit-maximizing firms self-select into costly voluntary pollution reduction programs (VPRs). Arora and Gangopadhy (1995) argue that firms want to attract a clientele of "green consumers" which are willing to pay more for goods produced in an environmentally friendly way. Voluntary pollution reductions may also deter lobbying by environmental groups for tighter environmental standards that "raise rivals' costs" (Innes and Bial, 2002); and avoid future environmental liability.

Empirical work has sought to determine the extent to which these theories have operated in practice to explain actual participation, most notably in the EPA's 33/50

[26] Joint work with Professor Robert Innes at The University of Arizona and Abdoul Sam at The Ohio State University.

program (e.g., see Arora and Cason, 1995, 1996; Videras and Alberini, 2000; Khanna and Damon, 1999; Sam and Innes, 2006). Empirical work has also documented the salutary effect of VPRs in reducing pollution (e.g., Khanna and Damon, 1999; Sam and Innes, 2006). However, to our knowledge, there has been no work to date identifying the mechanism by which VPRs may lead to emissions reductions, whether due to heightened management awareness and conscientiousness within given environmental systems and technologies or due to adoption of new environmental management systems (Anton, et al., 2004; Khanna, et al., 2005) or due to adoption of new environmental technologies.

In this paper, we study this last potential channel for beneficial effects of VPRs. In particular, VPRs could induce participant firms to innovate in their environmental technologies, thus lowering their costs of over-compliance. Pollutant reductions generally require costly reformulation of products or production processes, suggesting that over-compliance positively impacts environmental innovation. A limited number of theoretical (Milliman and Prince (1989), Fischer, and al. (2003)) and empirical (Jaffe and Palmer (1997), Brunnermeier and Cohen (2003), Carrión-Flores and Innes (2006)) papers study the impacts of various environmental policy instruments on technological innovation. While this work generally documents that higher pollutant abatement costs spur environmental R&D (the induced innovation hypothesis), none has explored the potential role of VPRs in promoting innovation.

We present an empirical study of the determinants of environmental innovation using a panel of 127 U.S. manufacturing industries defined by 3-digit SIC classifications over the 1989-2002 period. Following Brunnermeier and Cohen (2003) and Carrión-Flores and Innes (2006), we measure innovation by the number of successful environmental patent applications. The VPR that is our focus is the EPA's 33/50 program, the principal object of empirical work on VPRs to date. The 33/50 program was created in 1991 as the EPA's first formal effort to achieve voluntary pollution reductions by regulated firms. The program sought to reduce releases of seventeen toxic chemicals by a third by 1992 and by 50 percent by 1995, measured from 1988 baseline levels. At its inception, the program invited all firms releasing 33/50 pollutants in 1988 to participate (approximately 5000 companies). Although the 33/50 program was purely voluntary and its pollution reduction targets were not enforceable, there is ample anecdotal (EPA, 1999) and empirical evidence of the program's success. We measure 33/50 effects at a 3-digit SIC industry level using rates of industry participation in the program.

The reminder of the paper is organized as follows. In the next section we briefly describe the 33/50 program, followed by a description of competing empirical hypothesis that we test in this paper. In section 3.4, we describe our empirical model, followed by discussion of our data and econometric methods in section 3.5. Section 3.6 presents our estimation results. Finally, section 3.7 concludes.

3.2 The 33/50 Program

Started in 1991, the 33/50 program was the EPA's first formal effort to achieve voluntary pollution reductions by regulated firms. The program sought to reduce releases of seventeen toxic chemicals by a third by 1992 and by 50 percent by 1995, measured from 1988 baseline levels. Roughly seventy percent of the 33/50 chemicals (by 1988 weight of releases) were air pollutants (AC). Two of the chemicals (carbon tetrachloride and 1,1,1-trichloroethane) depleted the stratospheric ozone layer and, hence, came under the Montreal Protocol's provisions for the phase-out of such substances; however, these two chemicals represented less than fifteen percent of total 33/50 releases (in 1988).

The EPA initiated the 33/50 program shortly after creating the Toxic Release Inventory (TRI), a database compiling information on toxic releases of all firms with ten or more employees producing one or more of 320 targeted pollutants. In early 1991, the EPA invited the 509 companies emitting the largest volume of 33/50 pollutants to participate in the program; these companies were responsible for over three-quarters of total 33/50 releases as of 1988. In July 1991, the 4534 other companies with reported 33/50 releases in 1988 were asked to participate as well. With additional enrollments through 1995, the EPA invited a total of 10,167 firms to join the 33/50 program, and 1294 firms accepted. The latter program participants accounted for 58.8 percent of 33/50 releases in 1990.

The 33/50 program was purely voluntary and its pollution reduction targets were not enforceable. Despite the absence of apparent regulatory teeth, the EPA (1999) cites some aggregate statistics as indicators of the program's success. Among reporting firms, total 33/50 releases declined by over 52 percent between 1990 and 1996, and net 33/50 releases, excluding the two ozone-depleting compounds, declined by over 45 percent. In contrast, non-33/50 TRI releases fell by 25.3 percent over this period. Moreover, rates of 33/50 release reductions were greater for program participants (down 59.3 percent between 1990 and 1996) than for non-participants (down 42.9 percent over the same interval). Of course, these numbers may mask other hidden determinants of firms' pollution; for example, participating firms may have been more apt to reduce pollution, regardless of participation in the 33/50 program. However, recent work finds that, even controlling for other relevant explicators of pollution and potential selection bias in 33/50 program participation, the program led to significant reductions in participant emissions during the later program years of 1993-1995 (Sam and Innes, 2006; Khanna and Damon, 1999).

Conditional on the fact that participation in the 33/50 program, the question posed in this paper is whether participation in the 33/50 program spurred or impeded environmental research.

3.3 Empirical Hypotheses

Participation in the 33/50 program required a firm to file a plan which included a description of how it ptoposed to reduce its emissions of target pollutants. Moreover, the program conveyed some technical assistance to aid participants in accomplishing their target emission reductions. To the extent that firms take seriously any voluntary pollutant reduction commitments that they make (as is suggested by extant empirical evidence on the 33/50 program), such commitments may implicitly elevate the potential cost-reduction benefits of new environmental technologies, thus spurring more environmental R&D.

> ***Hypothesis 1.*** Higher rates of participation in the 33/50 program yield increased incentives for environmental R&D and, hence, more environmental patents.

Hypothesis 1 is essentially the “induced innovation” hypothesis as it applies to VPRs. The induced innovation hypothesis states that government policies affect research outcomes.

However, this logic may not reflect the true nature of the dynamic trade-offs that confront firms. For example, firms may have to decide how many resources to invest, alternately, in (i) environmental monitoring and compliance efforts that reduce short-run emissions, and (ii) research that may potentially yield new environmental compliance technologies. In principle, a VPR might spur a diversion of resources from the latter (R&D) to the former (monitoring and compliance). If so, we would instead have the competing hypothesis:

> ***Hypothesis 2.*** Higher rates of participation in the 33/50 program yield a redirection of resources away from environmental R&D and, hence, fewer environmental patents.

3.4 Empirical Model

Following Carrión-Flores and Innes (2006), we posit an underlying structural model that determines environmental patent outcomes as a function of anticipated emission standards, 33/50 participation rates, and other observable exogenous variables.

This model takes the following simple form:

$$P_{it} = a^*_{pit} + b^*_p E_{t-1}(Q_{it}) + c^*_p Q_{it-1} + d^*_p PR_{it-1} + f^*_p X_{pit} + \varepsilon^*_{pit} \tag{3.1}$$

where P_{it} is time t environmental patents in industry i, Q_{it} is time t emissions in industry i, $E_{t-1}(Q_{it})$ is its time $t-1$ expectation, PR_{it-1} is the appropriately lagged measure of the 33/50 participation rate in industry i, the X_{pit} represent exogenous variables, and ε^*_{pit} represent random errors. This structure gives rise to three types of joint endogeneity: (1) the observable regressor, Q_{it}, is jointly endogenous in the usual sense, with technological change potentially prompting revisions in emission standards; (2) the true regressor, $E_{t-1}(Q_{it})$, is measured with error; and (3) there is potential selection correlation between 33/50 participation rates and innovation because more innovative industries (ceteris paribus) may be more likely to participate in the VPR. Under perfect foresight, the second "endogeneity" problem evaporates and all we need are instruments that are highly correlated with emissions and 33/50 participation, but uncorrelated with patents. Without perfect foresight, but with rational expectations, we also need such identifying instruments, while also requiring that *all* instruments be lagged (see Carrión-Flores and Innes, 2006). In this paper, we estimate under both premises.

We use enforcement variables to jointly identify the two jointly endogenous variables (emissions and participation), as there is ample evidence that strict enforcement in the form of more inspections and enforcement actions, spurs emission reductions (Gray and, Deily 1996; LaPlante and Rilstone, 1995) and 33/50 participation (Sam and Innes, 2006), but has no discernable effect on innovation (Carrión-Flores and Innes, 2006; Brunnermeier and Cohen, 2003). Moreover, enforcement affects emissions which in turn affect innovation. However, we control for this mechanism, so no reason to expect direct effect on enforcement on innovation. Standard instrument and over-identifying restriction tests (Bound, et al., 1995; Wooldridge, 2002) provide statistical evidence in support of these instrument choices.

In estimating (3.1), there a number of determinants of innovative activity for which we need to control, in addition to the evident role of anticipated emissions and the key 33/50 participation effects of interest. First, spillovers in research activity may lead to environmental patent successes; we use non-environmental patent counts and overall industry R&D investments to control for these effects. Second – and unlike prior work – we seek to measure potential impacts of environmental pressure groups on environmental R&D. We do so by including an industry measure of Sierra Club strength. Third, we control for industry size using a measure of real sales. Fourth, we control for the nature of industry assets, both age and capital intensity. We expect industries with older assets and more capital intensive production to have more scope for cost-reducing environmental R&D. Fifth, more rapidly growing industries may have either more or less incentive to innovate, whether because they have already modernized and hence have less scope for innovation or because they are more innovative by nature and hence more likely to innovate in the environmental realm as well. Finally, more concentrated industries may be prone to either more or less environmental R&D. On one hand, concentration gives rise to "raising rivals costs" motives for heightened research (Innes and Bial, 2002); on the other, concentrated industries may collectively recognize the costs of higher R&D in spurring tightened environmental regulations, leading to a research deterrent (Carrión-Flores and Innes, 2006). We control for such effects by including a measure of industry concentration in our estimations.

So far, in deriving our patent equation (3.1), we have implicitly assumed a linear process that generates a continuous variable. However, measured patent outcomes take a count form, with no negative values, a substantial number of zeroes (roughly one third in our sample), and integer positive values that range from one to 153 (with half of the positive values less than 40). Conceptually, we interpret patent outcomes as the observable consequence of our continuous (and unobservable) index of technology change P_{it} (of equation (3.1)). Specifically, let us suppose that patent counts P^*_{it} are distributed Poisson with

$$E(P^*_{it} \mid \varepsilon^*_{pit}) = \exp(P_{\mathrm{it}}), \tag{3.2}$$

where P_{it} is determined by equation (3.2) with industry fixed effects.[27] This gives us the multiplicative error Poisson panel model, with endogenous regressors, of Blundell, Griffith and Windmeijer (2002) (see also Windmeijer (2002) and Windmeijer and Santos Silva (1997)). This is the model we use to estimate the patent equation.[28]

3.5 Data Description and Definition of Variables

Our sample is an unbalanced panel of 127 manufacturing industries defined by 3-digit SIC classifications (SIC codes 200-399) over the period 1989 – 2002. Total toxic emissions data are available from the EPA's Toxic Release Inventory (TRI) for 1989-2002.[29] Using the TRI, we construct industry level total toxic releases (***Tri-Rel***) by aggregated weight by year. Facility releases reported in the TRI are assigned to the primary industry of the parent company.

Following previous studies (c.f., Jaffe and Palmer, 1997; Brunnermeier and Cohen, 2003 and Carrión-Flores and Innes, 2006), we use successful environmental patent applications as a proxy for environmental innovation. Using data from the U.S. Patent

---

[27] As in the emission equation, we allow for both time and industry fixed effects. However, the time dummies are again jointly insignificant; hence, for efficiency, we estimate with industry fixed effects only.

[28] Because we have a mixture Poisson with multiplicative error, our estimation allows for over-dispersion (see Cameron and Trivedi, 1998, p. 98) and thus avoids the main criticism of a standard fixed effects Poisson.

[29] Because the first year of TRI release reports are considered incomplete and suspect, we only rely on post-1988 TRI data.

and Trademark Office, we construct successful patent application counts by year, by industry, environmental and non-environmental, obtained by U.S. companies. Environmental patents are determined by patent classifications that relate to air or water pollution, hazardous waste prevention, disposal and control, recycling and alternative energy (***Env-Pat***). Non-environmental patents are those in all other patent utility classes (***Nonenv-Pat***). For robustness purposes, we also use a narrower measure of environmental patent count based on the categorization of Brunnermeier and Cohen (2003); we denote this measure (***Env-PatBC***).

The EPA's Office of Environmental Information Records provided data on 33/50 participation, as well as Federal and State enforcement activity under the Clean Air Act (CAA).[30] We measure 33/50 effects at a 3-digit SIC industry level using rates of industry participation in the program. Specifically, for each year of the 33/50 program (1991-1995), we measure the industry participation rate (***PR***) as the proportion of reported industry 33/50 emissions attributable to program participants in that year.

From our enforcement data, we construct three measures of enforcement stringency: (1) counts of Federal and State enforcement actions (***Actions***), (2) the numbers of facilities out of compliance with clean air laws (***Outcomp***), and (3) the numbers of reported self-inspections (***Selfinspect***). We use these enforcement variables to jointly identify emissions and participation rates (with three lags due to plausible lags in effects and because such lags avoid any potential for joint endogeneity).

The Sierra Club provided us a panel of its state membership. To obtain a measure of

[30] Because the 33/50 program primarily relates to air releases, we focus on enforcement activity under the CAA.

environmental group influence on each industry, we obtain the weighted average annual Sierra Club membership for the industry (***Sierra***), weighting each state's membership by the proportion of the industry's regulated facilities operating in the state.

Financial and employment data was obtained from the Standard and Poor's Compustat Dataset. Deflators are obtained using producer price indexes reported in the Economic Report of the President (2004). For controls, we include (deflated) industry sales volume (***Sales***) in order to account for potential effects of industry size on patents; a measure of capital intensity (***Cap-Int***), namely, the level of new capital and equipment expenditures divided by sales volume; the industry's total lagged level of research and development expenditures per-unit-sales (***R&D***) in order to capture effects of overall industry research activity on environmental innovation; a measure of asset age (***Age)***, obtained by dividing total assets of an industry by its gross assets (as in Khanna and Damon, 1999); and and environmental policy may be affected by the rates of growth, and hence the modernity, of the different industries. We therefore include a sales growth measure (***Salesgr)***. To measure industry concentration, we construct a four-firm Herfindahl index (***Concent***) using annual sales data reported in Compustat. [31]

Merging all of our datasets yielded 127 industries for the 1989-2002 period for a total of 1778 observations over the 14 year period. If we incorporate two year lags in emissions, we have an unbalanced panel of 127 industries over 1991-2002 for a total of 1397 observations. Summary statistics are presented in Table 3.1.

[31] Initially, we also included an exogenous variable measuring each industry's export intensity (ratio of exports to total sales). However, as this variable was not statistically significant in any estimated equation (regardless of the model), and its inclusion compromised model performance, we do not include it in our reported model estimations.

Variable Name	Mean	Std. Dev.
TRI-REL	39.52	145.1
SIERRA	79.54	94.08
SALES	31115	103557
SALESGR	-0.0347	0.2642
AGE	0.7146	0.1430
CAPINT	0.0824	0.0513
R&D INTENSITY	0.5638	0.284
CONCENTRATION	0.0985	0.2298
ENV-PATENTS	21.28	17.45
ENV-PATENTSBC	8.93	12.28
NONENV-PATENTS	26.69	11.45
PR (1991-1995)	0.4940	0.2925
Two-Lags / Most Recent		
PR2002	0.4907	0.2429
PR2001	0.4786	0.2423
PR2000	0.4562	0.2378
PR1999	0.4503	0.2445
PR1998	0.4761	0.2542
PR1997	0.4632	0.2664
PR1996	0.4686	0.2687
PR1995	0.4928	0.2572
PR1994	0.4713	0.2376
PR1993	0.4995	0.2758
Instruments		
ACTIONS	86.34	169.63
OUTCOMP	112.81	178.34
SELFINSPECT	5.17	13.43

TABLE 3.1. Summary Statistics

Our measure of innovation (number of patents) takes a count form with a relatively large number of zero's and a mean count of approximately 20. As noted above, we also have endogeneity issues with respect to both emissions and participation rates. With regard to emissions, we measure the "true regressor," $E_{t-1}(Q_t)$, using actual emissions. Due to both joint endogeneity and measurement error, we instrument in order to avoid bias and inconsistency. Due to serial correlation in our emissions data, endogeneity bias also attaches to our lagged emission regressor; hence, we also instrument this variable using appropriately lagged exogenous data.

To estimate our count panel model with endogenous regressors, we follow Windmeijer and Silva (1997) and estimate a Poisson fixed effects structure by the Generalized Method of Moments (GMM). Testing for fixed time effects, we find none of significance and hence estimate without.

Effects of 33/50 participation on patent applications occur at least with the lag required for research to produce patentable outcomes (which we assume to be at least two years). Indeed, these effects may potentially take several years to manifest themselves. We attempt to capture lag-specific effects of participation by constructing year-specific participation rate variables. With a minimum of two-year lags in potential effects, we measure participation impacts in each of years 1993 (two years after the 33/50 program's inception) and 1997 by constructing year-specific two-year-lagged participation rate variables. For years 1998-2002, we measure participation impacts with year-specific variables for the most recent (1995) participation rates. These variables permit us to estimate year-specific participation effects (with variables $\boldsymbol{PR_t}$, t=1993,…2002).[32]

Turning next to instruments that identify emissions and participation rates (in the patent equation); we note that environmental enforcement activity is widely cited as a stimulus to pollution abatement (e.g., see Magat and Viscusi (1990), Gray and Deily (1996), Deily and Gray (2007), Decker and Pope (2006)). However, there is no evidence, in theory or empirical work, that enforcement activity affects innovative activity other than due to its effects on "effective" environmental standards and, hence, emissions.[33] We therefore use various measures of U.S. environmental enforcement activity to identify emissions. Specifically, environmental compliance and enforcement histories are obtained from the EPA's IDEA database. IDEA contains facility level data from the Aerometric Information Retrieval System (AIRS) and the Air Facility Subsystem (AFS). AFS contains compliance and enforcement data on stationary sources of air pollution. Regulated sources range from large industrial facilities to relatively small operations. We use counts of enforcement actions (***Actions***), numbers of facilities out of compliance with clean air laws (***Outcomp***), and the number of reported self-inspections (***Selfinspect***) as indicators of environmental enforcement stringency. Because enforcement effects on emission performance occur with a substantial delay, we lag all of our instruments by

---

[32] To identify the year-specific participation variables, we construct year-specific counterparts to our lagged enforcement variables.

[33] Brunnermeier and Cohen (2003) include a measure of government environmental inspections as an explanatory variable in their patent equation. In doing so, they rightfully argue (p. 284) that "to the extent that stricter government monitoring or enforcement induces firms to comply, they might now seek less costly methods of complying." In our model, in contrast, compliance efforts (that may spur innovation) are captured by our emissions measure; that is, compliance efforts will reduce emissions, which in turn will potentially fuel environmental R&D incentives. In sum, in our paper, enforcement effects operate via emissions, even though they need not operate via PAE, the policy proxy in Brunnermeier and Cohen's (2003) analysis.

three years.[34] For robustness purposes, we consider a variety of different instrument combinations; we report results using two combinations but have obtained similar results using other instruments menus.

3.6 Empirical Estimation and Findings

To judge the strength of our identifying (enforcement) instruments, we first estimate relevant "first stage" models of emissions and participation rates (per standard practice, Bound, et al., 1995). Because emissions are continuous, the relevant first-stage is a fixed effects linear panel estimation with all exogenous data, as reported in Table 3.2; as indicated, the three enforcement variables are highly correlated with emissions, with predicted signs. In particular, enforcement actions and heightened regulatory scrutiny due to out-of-compliance status both spur emission reductions; conversely, self-inspections – by potentially preempting regulatory scrutiny – are associated with higher emissions.

---

[34] Lagging three years has the added advantage of avoiding any potential for endogeneity between emissions and enforcement. We considered other enforcement lags and found that three-year lags in our three enforcement variables performed the best as determinants of emissions.

	Model 1: Fixed Effects		Model 2: Dynamic Model	
Dependent Variable	***Emissions***			
Variable Instrumented	***None***		***Emissions*** $_{t-2}$	
Variables	Coefficient (Robust SE)	z-ratio	Coefficient (Robust SE)	z-ratio
SELFINSPECT $_{t-3}$	17.13 (4.78)	3.58	13.33 (3.98)	3.34
OUTCOMP $_{t-3}$	-195.78 (71.78)	-2.72	-65.16 (36.24)	-1.80
ACTIONS $_{t-3}$	-56.37 (11.29)	-4.99	-56.16 (36.60)	-1.53
R&D EXPENDITURES $_{t-2}$	-0.1205 (0.3293)	-0.36	-0.2266 (0.1145)	-1.98
CAPITAL INTENSITY	1.0293 (0.7832)	1.31	11.60 (34.97)	0.33
CONCENTRATION	-102.02 (76.11)	-1.34	-80.27 (11.25)	-7.13
AGE	50.90 (39.82)	1.27	73.90 (32.82)	2.25
SALES	0.0542 (0.0713)	0.76	0.0742 (0.0513)	1.45
SALES GROWTH	-3.73 (4.29)	-0.86	-11.69 (3.51)	-3.33
USNONAPP $_{t-2}$	19.29 (4.74)	4.06	14.82 (10.36)	1.43
SIERRA	13.39 (10.28)	1.30	25.83 (16.99)	1.52
EMISSIONS $_{t-1}$	*	*	0.7201 (0.0532)	13.53
CONSTANT	-1495.33 (1044.28)	-1.43	-21.07 (18.62)	-1.13
	F-Statistic	p-value	Chi-Statistic	p-value
Instrument Tests	7.59	0.0001	13.27	0.0051
Time Dummy Tests	1.15	0.3295	0.59	0.8521
Hansen Tests	*	*	39.29	0.472

TABLE 3.2. Emission Equation Fixed effects (First Stage)

For our 33/50 participation rate data, the appropriate "first stage" procedure is complicated by the non-continuous and truncated nature of the data (with all rates in the unit interval and a number of zeros). To account for this data structure, we perform a "first stage" Tobit estimation using the participation odds ratios ($\boldsymbol{PR}/(1-\boldsymbol{PR})$) (which are truncated at zero and otherwise continuous on R^+). Results are reported in Table 3.3. As required, the enforcement instruments are highly correlated with our participation variable (see, for example, the chi-square test statistic for the null of zero coefficients on our three enforcement variables).

	Model 1: Random Effects Tobit		Model 2: Random Effects Tobit	
Dependent Variable	***PR***			
Instruments Lags	***t-2***		***t-3***	
Variables	Coefficient (Robust SE)	z-ratio	Coefficient (Robust SE)	z-ratio
SELFINSPECT	0.0139 (0.0037)	1.89	0.0136 (0.0040)	3.38
OUTCOMP	0.0504 (0.0120)	4.20	0.0304 (0.0104)	2.92
ACTIONS	-0.0517 (0.0160)	-3.22	-0.0489 (0.0161)	-3.03
R&D EXPENDITURES $_{t-2}$	0.0001 (0.00008)	1.18	0.0001 (0.00007)	1.36
CAPITAL INTENSITY	21.4145 (14.2628)	1.50	10.4380 (15.3953)	0.68
CONCENTRATION	31.4405 (4.1744)	7.53	33.8602 (4.1159)	8.23
AGE	-8.4483 (6.1804)	-1.37	-10.4885 (5.3931)	-1.94
SALES	-0.00009 (0.00005)	-1.78	-0.00008 (0.00004)	-1.87
SALES GROWTH	-1.0027 (2.5886)	-0.39	-1.0652 (2.6114)	-0.41
USNONAPP $_{t-2}$	-0.0160 (0.0030)	-5.31	-0.0148 (0.0030)	-4.95
SIERRA	-0.0058 (0.0100)	-0.58	0.0003 (0.0089)	0.04
CONSTANT	-24.820 (4.5812)	-5.42	-24.057 (4.1895)	-5.74

TABLE 3.3. Participation Rate Equation (First Stage)

Tables 3.4-3.6 report estimations of our perfect foresight model. We use two types of lags in the participation rate variable. The first one is a two year lag in participation rate data for the early post program years (1993-1997) and the most recent lag for the later period (1998-2002); however, because 33/50 participation may have effects on research with more than a two year lag, we use three-year lags in participation rate data (rather than two) for the early post-program years (1994-1998) and the most recent lag for the later period (1999-2002). Table 3.4 includes estimation of year-specific participation rate effects. Table 3.5 Model 3 includes estimations that group participation rate effects for the earlier post-program years (1993-1997) and the later years (1998-2002). Model 4 includes the estimation using a three year lag in the participation rate for the earlier post-program years (1994-1998) and the later years (1999-2002). Finally, Table 3.6 includes the estimations that again group participation rate effects for the entire sample period using two year lags (1993-2002) and three year lags (1994-2002).

Several results should be stressed. First and foremost, in all models, we find significant negative long-run effects of 33/50 participation on successful environmental patents, broadly supporting Hypothesis 2. In Table 3.4 model 1, for example, we have statistically significant negative effects of 33/50 participation rates on successful patents in years 1993, 1995, 2000, and 2002.

Second, as in Carrión-Flores and Innes (2006), the anticipation of tightened emission standards spurs more environmental R&D (with significant negative coefficients on ***Tri-Rel***. Third, environmental innovation is positively associated with industry scale (***Sales***), research spillovers (***R&D*** and ***Nonenv-Pat***), capital intensity (***Capint***), and older assets (with ***Age*** measuring "newness" of assets), all as expected. Concentration (***Concent***) is found to decrease innovation, consistent with the view that concentrated industries recognize regulatory responses to improved technologies and circumscribe their R&D investments accordingly. More rapidly growing industries are also found to be the source of less environmental innovation (with negative effects of the ***Salesgr*** variable), perhaps because they are already quite modern and hence lesser scope for technological improvement.

Finally, we should also stress our test of instrument performance. For all models, the implicit (Sargan) test of zero correlation between our identifying instruments and the equation error is not rejected at reasonable levels of significance.

	Model 1: Perfect Foresight		Model 2: Perfect Foresight	
Dependent Variable	*Env. Patents*			
Variable Instrumented	*TRI-REL* $_{t-2}$, *PR*		*TRI-REL* $_{t-2}$, *PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0106 (0.0037)	0.0043	-0.0106 (0.0194)	0.5861
TRI-REL $_{t-2}$	-0.0330 (0.0154)	0.0326	-0.0125 (0.0115)	0.2758
SIERRA	-0.0022 (0.0005)	0.0000	-0.0021 (0.0040)	0.5936
SALES	0.0132 (0.0022)	0.0000	0.0149 (0.0144)	0.3029
SALES GROWTH	-0.0143 (0.0684)	0.8344	-0.0265 (0.3758)	0.9438
AGE	-1.4501 (0.4665)	0.0019	0.5490 (1.7498)	0.7537
CAPITAL INTENSITY	2.0257 (0.4286)	0.0000	4.2919 (6.4653)	0.5068
R&D EXPENDITURES $_{t-2}$	0.0322 (0.2102)	0.8744	0.6972 (1.5344)	0.6495
USNONAPP $_{t-2}$	0.0121 (0.0003)	0.0000	0.0143 (0.0046)	0.0018
CONCENTRATION	-4.8328 (1.1567)	0.0000	-7.4096 (7.8057)	0.3425
PR1993	-4.9010 (1.3452)	0.0003	*	*
PR1994	0.4369 (0.8205)	0.5944	2.2129 (5.5195)	0.8651
PR1995	-0.7409 (0.2062)	0.0003	-2.7913 (1.8501)	0.1314
PR1996	-0.4993 (0.4077)	0.2207	-0.1364 (1.5374)	0.9293
PR1997	-0.0538 (0.9456)	0.9546	0.5972 (2.9479)	0.8395
PR1998	0.1191 (0.6480)	0.8541	5.2582 (7.8704)	0.5041
PR1999	-0.3110 (1.5210)	0.8380	2.7090 (4.5350)	0.5503
PR2000	-4.6646 (0.6230)	0.0000	-3.1218 (3.6590)	0.3935
PR2001	-7.7300 (1.0168)	0.0000	-2.8887 (3.4861)	0.4073
PR2002	-5.9213 (1.1000)	0.0000	-0.4893 (5.6165)	0.9306
	Statistic	p-value	Statistic	p-value
Hansen Tests	20.1475	0.2668	13.7571	0.0325

TABLE 3.4. Perfect Foresight Estimation Results

	Model 1: Perfect Foresight		Model 2: Perfect Foresight	
Dependent Variable	*Env. Patents*			
Variable Instrumented	*TRI-REL* $_{t-2}$,*PR*		*TRI-REL* $_{t-2}$,*PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0029 (0.0097)	0.7656	-0.0009 (0.1509)	0.8800
TRI-REL $_{t-2}$	-0.1387 (0.1569)	0.3767	-0.0455 (0.0105)	0.0000
SIERRA	-0.0032 (0.0030)	0.2800	-0.0012 (0.0003)	0.0000
SALES	0.0078 (0.0049)	0.1078	0.0133 (0.0044)	0.0023
SALES GROWTH	-1.1602 (0.2587)	0.0000	-0.0324 (0.0652)	0.6190
AGE	0.8847 (2.5295)	0.7265	-1.4716 (0.2197)	0.0000
CAPITAL INTENSITY	4.6932 (3.9216)	0.2314	1.5109 (0.5150)	0.0033
R&D EXPENDITURES $_{t-2}$	0.3021 (1.6419)	0.8540	0.4990 (0.2321)	0.0316
USNONAPP $_{t-2}$	0.0132 ().0024)	0.0000	0.0117 (0.0004)	0.0000
CONCENTRATION	-2.8634 (2.2227)	0.4248	-4.8180 (1.1576)	0.0000
PR1999-02	-1.8114 (2.2697)	0.4248	*	*
PR1994-98	-0.8909 (1.5740)	0.5714	*	*
PR1993-97	*	*	-0.5961 (0.2170)	0.0060
PR1998-02	*	*	-1.5290 (1.0295)	0.1375
	Statistic	p-value	Statistic	p-value
Hansen Tests	11.5429	0.1166	19.7024	0.3498

TABLE 3.5. Perfect Foresight Estimation Results

	Model 3: Perfect Foresight		Model 4: Perfect Foresight	
Dependent Variable	*Env. Patents*			
Variable Instrumented	*TRI-REL* $_{t-2}$, *PR*		*TRI-REL* $_{t-2}$, *PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0030 (0.0062)	0.6209	-0.0031 (0.0079)	0.6979
TRI-REL $_{t-2}$	-0.0364 (0.0113)	0.0013	-0.0179 (0.0058)	0.0021
SIERRA	-0.0012 (0.0002)	0.0000	-0.0011 (0.0021)	0.5913
SALES	0.0124 (0.0042)	0.0035	0.0080 (0.0096)	0.4093
SALES GROWTH	-0.0484 (0.0451)	0.2836	0.0876 (0.5156)	0.8651
AGE	-1.3959 (0.1895)	0.0000	1.5184 (2.6108)	0.5609
CAPITAL INTENSITY	1.4709 (0.3146)	0.0000	12.7432 (7.2064)	0.0770
R&D EXPENDITURES $_{t-2}$	0.5418 (0.2508)	0.0308	1.0698 (0.9920)	0.2809
USNONAPP $_{t-2}$	0.0109 (0.0005)	0.000	0.0143 (0.0023)	0.0000
CONCENTRATION	-4.2402 (1.2439)	0.0007	-3.4596 (1.2409)	0.0053
PR1993-02	-0.5892 (0.2211)	0.0077	*	*
PR1994-02	*	*	-1.4902 (0.6089)	0.0144
	Statistic	p-value	Statistic	p-value
Hansen Tests	18.9565	0.3945	9.3075	0.2313

TABLE 3.6. Perfect Foresight Estimation Results

Tables 3.7-3.9 report the estimations for the rational expectations model; we follow the same year specific and grouped participation rate effects as the perfect foresight model estimations. The main difference is that the exogenous variables are lagged two periods. We find significant negative long-run effects of 33/50 participation on successful environmental patents for the group participation rate effects for the entire sample period with two year lags (1993-2002) and three year lags (1994-2002). This result is consistent when we divide the participation rate into short-run effects (1993-1997) and long-run effects (1998-2002). Moreover, in Table 3.9, for example, we have statistically significant negative effects of 33/50 participation rates on successful patents in years 1993, 1996, 2000, 2001, and 2002.

We also obtain negative and significant coefficients on ***Tri-Rel*** which means that the anticipation of tightened emission standards spurs more environmental R&D. Third, we find that environmental innovation is positively associated with industry. Concentration (***Concent***) is found to decrease innovation, consistent with the view that concentrated industries recognize regulatory responses to improved technologies and circumscribe their R&D investments accordingly. We also find that more capital intensive industries with older assets may have more scope and incentive for emission-reducing innovation. Finally, more rapidly growing industries are also found to be the source of less environmental innovation (with negative effects of the ***Salesgr*** variable), perhaps because they are already quite modern and hence lesser scope for technological improvement.

	Model 1: Rational Expectations		Model 2: Rational Expectations	
Dependent Variable	*Env. Patents*			
Variable Instrumented	*TRI-REL* $_{t-2}$, *PR*		*TRI-REL* $_{t-2}$, *PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0015 (0.0042)	0.7141	-0.0021 (0.0033)	0.5250
TRI-REL $_{t-2}$	-0.0876 (0.0109)	0.0000	-0.0855 (0.0109)	0.0000
SIERRA $_{t-2}$	-0.0007 (0.0003)	0.0074	-0.0007 (0.0003)	0.0084
SALES $_{t-2}$	0.0120 (0.0021)	0.0000	0.0119 (0.0021)	0.0000
SALES GROWTH $_{t-2}$	0.2869 (0.0528)	0.0000	0.2867 (0.0559)	0.0000
AGE $_{t-2}$	-2.3208 (0.3560)	0.0000	-2.3131 (0.3681)	0.0000
CAPITAL INTENSITY $_{t-2}$	1.5459 (0.5431)	0.0044	1.5024 (0.5374)	0.0052
R&D EXPENDITURES $_{t-2}$	-0.7518 (0.1371)	0.0000	-0.7469 (0.1392)	0.0000
USNONAPP $_{t-2}$	0.0115 (0.0003)	0.0000	0.0115 (0.0003)	0.0000
CONCENTRATION $_{t-2}$	-3.9564 (1.1064)	0.0003	-4.0252 (1.0727)	0.0002
PR1994-02	*	*	-0.3261 (0.0798)	0.0000
PR1993-02	-0.3447 (0.0803)	0.0000	*	*
	Statistic	p-value	Statistic	p-value
Hansen Tests	20.2590	0.3184	20.5906	0.3006

TABLE 3.7. Rational Expectations Estimation Results

	Model 3: Rational Expectations		Model 4: Rational Expectations	
Dependent Variable	*Env. Patents*			
Variable Instrumented	*TRI-REL* $_{t-2}$, *PR*		*TRI-REL* $_{t-2}$, *PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0035 (0.0078)	0.6549	-0.0033 (0.0031)	0.2888
TRI-REL $_{t-2}$	-0.0867 (0.0116)	0.0000	-0.0903 (0.0139)	0.0000
SIERRA $_{t-2}$	-0.0007 (0.0003)	0.0051	-0.0005 (0.0003)	0.1114
SALES $_{t-2}$	0.0098 (0.0016)	0.0000	0.0118 (0.0019)	0.0000
SALES GROWTH $_{t-2}$	0.3408 (0.0767)	0.0000	0.3341 (0.0688)	0.0000
AGE $_{t-2}$	-2.7258 (0.4997)	0.0000	-2.5432 (0.4323)	0.0000
CAPITAL INTENSITY $_{t-2}$	2.2706 (0.7697)	0.0032	1.4678 (0.5826)	0.0118
R&D EXPENDITURES $_{t-2}$	-0.6830 (0.1209)	0.0000	-0.7977 (0.1306)	0.0000
USNONAPP $_{t-2}$	0.0106 (0.0003)	0.0000	0.0116 (0.0003)	0.0000
CONCENTRATION $_{t-2}$	-3.6840 (1.0931)	0.0008	-4.2522 (1.0845)	0.0001
PR1993-97	-0.2671 (0.0704)	0.0001	*	*
PR1998-02	-1.1520 (0.4965)	0.0203	*	*
PR1994-98	*	*	-0.3438 (0.0711)	0.0000
PR1998-02	*	*	-1.5865 (0.9402)	0.0915
	Statistic	p-value	Statistic	p-value
Hansen Tests	23.5443	0.1705	22.7082	0.2021

TABLE 3.8. Rational Expectations Estimation Results

	Model 5: Rational Expectations		Model 6: Rational Expectations	
Dependent Variable	*Env. Patents*			
Variable Instrumented	*TRI-REL* $_{t-2}$*,PR*		*TRI-REL* $_{t-2}$*,PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0245 (0.0100)	0.0145	-0.0305 (0.0140)	0.0290
TRI-REL $_{t-2}$	-0.0367 (0.0155)	0.0179	-0.0380 (0.0088)	0.0000
SIERRA $_{t-2}$	-0.0023 (0.0003)	0.0000	-0.0019 (0.0003)	0.0000
SALES $_{t-2}$	0.0140 (0.0029)	0.0000	0.0142 (0.0023)	0.0000
SALES GROWTH $_{t-2}$	0.0677 (0.0718)	0.3461	0.0692 (0.0617)	0.2626
AGE $_{t-2}$	-2.0354 (0.5375)	0.0002	-2.6739 (0.7662)	0.0005
CAPITAL INTENSITY $_{t-2}$	2.3781 (0.9397)	0.0114	2.7242 (0.7944)	0.0006
R&D EXPENDITURES $_{t-2}$	-0.7561 (0.1481)	0.0000	-0.9025 (0.1631)	0.0000
USNONAPP $_{t-2}$	0.0117 (0.0005)	0.0000	0.0113 (0.0004)	0.0000
CONCENTRATION $_{t-2}$	-4.3194 (1.3258)	0.0011	-4.7598 (1.3338)	0.0003
PR1993	-5.3220 (1.4413)	0.0002	*	*
PR1994	0.1270 (0.5398)	0.8140	6.6395 (7.8013)	0.3947
PR1995	-0.2460 (0.1638)	0.1332	-0.1900 (1.3291)	0.8863
PR1996	-0.2412 (0.1347)	0.0733	0.1019 (0.0741)	0.1690
PR1997	0.2639 (0.7920)	0.7390	0.2506 (0.7014)	0.7208
PR1998	0.0615 (0.7564)	0.9352	0.1758 (0.7575)	0.8164
PR1999	0.3973 (1.8886)	0.8334	-0.1809 (1.4015)	0.8973
PR2000	-5.7115 (0.7839)	0.0000	-5.2718 (1.1993)	0.0000
PR2001	-8.1318 (1.0345)	0.0000	-7.4406 (0.6421)	0.0000
PR2002	-7.2718 (1.3111)	0.0000	-6.4393 (1.1159)	0.0000
	Statistic	p-value	Statistic	p-value
Hansen Tests	19.7194	0.2888	23.3376	0.1386

TABLE 3.9. Rational Expectations Estimation Results

Tables 3.10-3.12 report estimations for the rational expectation model using Brunnermeier and Cohen (2003) environmental patent counts categorization. The purpose of estimating with a narrower environmentla patent count is to show robustness in our estimations. From Table 3.10, we note negative and statistically significant long-run effects of 33/50 participation on successful environmental patents for the group participation rate effects for the entire sample period with two year lags (1993-2002) and three year lags (1994-2002).

	Model 1: Rational Expectations		Model 2: Rational Expectations	
Dependent Variable	*Env. Patents BC*			
Variable Instrumented	*TRI-REL* $_{t-2}$, *PR*		*TRI-REL* $_{t-2}$, *PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0005 (0.0083)	0.9544	-0.0004 (0.0083)	0.9631
TRI-REL $_{t-2}$	-0.0581 (0.0056)	0.0000	-0.0800 (0.0160)	0.0000
SIERRA $_{t-2}$	-0.0001 (0.0002)	0.7317	-0.0099 (0.0005)	0.0977
SALES $_{t-2}$	-0.0103 (0.0030)	0.0007	0.0099 (0.0020)	0.0000
SALES GROWTH $_{t-2}$	0.3123 (0.0732)	0.0000	0.1699 (0.1372)	0.2159
AGE $_{t-2}$	-2.5600 (0.3775)	0.0000	-2.0909 (0.2948)	0.0000
CAPITAL INTENSITY $_{t-2}$	2.9895 (0.4465)	0.0000	1.7855 (0.6047)	0.0032
R&D EXPENDITURES $_{t-2}$	-1.3271 (0.0710)	0.0000	-1.4005 (0.0770)	0.0000
USNONAPP $_{t-2}$	0.0113 (0.0005)	0.0000	0.0112 (0.0003)	0.0000
CONCENTRATION $_{t-2}$	-3.4462 (1.2752)	0.0069	-3.3107 (1.0801)	0.0000
PR1994-02	*	*	-0.0128 (0.0107)	0.0005
PR1993-02	-0.2461 (0.0805)	0.0022	*	*
	Statistic	p-value	Statistic	p-value
Hansen Tests	19.4632	0.3638	24.4839	0.1398

TABLE 3.10. Rational Expectations Estimation Results using Env. BC Counts

	Model 3: Rational Expectations		Model 4: Rational Expectations	
Dependent Variable	*Env. Patents BC*			
Variable Instrumented	*TRI-REL* $_{t\text{-}2}$, *PR*		*TRI-REL* $_{t\text{-}2}$, *PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0009 (0.0077)	0.9049	-0.0040 (0.0099)	0.6816
TRI-REL $_{t\text{-}2}$	-0.0557 (0.0065)	0.0000	-0.0714 (0.0099)	0.0000
SIERRA $_{t\text{-}2}$	-0.0001 (0.0003)	0.6376	-0.0003 (0.0002)	0.2473
SALES $_{t\text{-}2}$	0.0099 (0.0022)	0.0000	0.0105 (0.0028)	0.0002
SALES GROWTH $_{t\text{-}2}$	0.1497 (0.1081)	0.1663	0.2994 (0.0870)	0.0006
AGE $_{t\text{-}2}$	-2.9979 (0.4109)	0.0000	-2.9537 (0.4064)	0.0000
CAPITAL INTENSITY $_{t\text{-}2}$	4.0897 (0.7151)	0.0000	3.0228 (0.4313)	0.0000
R&D EXPENDITURES $_{t\text{-}2}$	-1.3434 (0.0726)	0.0000	-1.3410 (0.0770)	0.0000
USNONAPP $_{t\text{-}2}$	0.0115 (0.0006)	0.0000	0.0115 (0.0005)	0.0000
CONCENTRATION $_{t\text{-}2}$	-2.9935 (1.1322)	0.0082	-3.8399 (1.3109)	0.0034
PR1993-97	-0.0749 (0.0708)	0.2904	*	*
PR1998-02	-1.2508 (0.5008)	0.0125	*	*
PR1994-98	*	*	-0.0987 (0.0495)	0.0461
PR1998-02	*	*	-1.1540 (0.9669)	0.2327
	Statistic	p-value	Statistic	p-value
Hansen Tests	15.8987	0.5996	18.6073	0.4164

TABLE 3.11. Rational Expectations Estimation Results using Env. BC Counts

	Model 5: Rational Expectations		Model 6: Rational Expectations	
Dependent Variable	*Env. Patents BC*			
Variable Instrumented	*TRI-REL* $_{t-2}$*,PR*		*TRI-REL* $_{t-2}$*,PR*	
Variables	Coefficient (Robust SE)	p-value	Coefficient (Robust SE)	p-value
TRI-REL	-0.0026 (0.0031)	0.4092	-0.0030 (0.0058)	0.6041
TRI-REL $_{t-2}$	-0.0841 (0.0096)	0.0000	-0.0635 (0.0057)	0.0000
SIERRA $_{t-2}$	-0.0003 (0.0004)	0.9432	-0.0006 (0.0004)	0.0866
SALES $_{t-2}$	0.0166 (0.0021)	0.0000	0.0114 (0.0030)	0.0002
SALES GROWTH $_{t-2}$	0.2928 (0.1315)	0.0260	0.0117 (0.0753)	0.8761
AGE $_{t-2}$	-1.7403 (0.5751)	0.0025	-1.4958 (0.5048)	0.0030
CAPITAL INTENSITY $_{t-2}$	3.6425 (0.9819)	0.0002	3.9981 (0.4524)	0.0000
R&D EXPENDITURES $_{t-2}$	-1.4573 (0.0794)	0.0000	-1.6256 (0.0864)	0.0000
USNONAPP $_{t-2}$	0.0107 (0.004)	0.0000	0.0119 (0.0004)	0.0000
CONCENTRATION $_{t-2}$	-4.5009 (1.2029)	0.0002	-3.2390 (1.3184)	0.0140
PR1993	-2.2467 (1.0193)	0.0275	*	*
PR1994	1.5038 (1.6834)	0.3717	1.6075 (3.7205)	0.6657
PR1995	-0.0049 (0.0677)	0.9425	-0.2088 (1.9956)	0.9167
PR1996	-1.4699 (0.4851)	0.0024	-0.0037 (0.0745)	0.9600
PR1997	2.1921 (1.0985)	0.0460	3.0300 (2.4940)	0.2244
PR1998	1.0273 (0.6889)	0.1359	0.4389 (0.7543)	0.5606
PR1999	-2.1588 (1.6813)	0.1991	-0.8044 (1.6646)	0.6289
PR2000	-7.4420 (0.4073)	0.0000	-3.3121 (1.3481)	0.0140
PR2001	-1.9053 (5.5834)	0.7329	-7.6407 (0.7093)	0.0000
PR2002	-8.7256 (0.5022)	0.0000	-6.8618 (0.9146)	0.0000
	Statistic	p-value	Statistic	p-value
Hansen Tests	15.8673	0.5333	21.2233	0.2165

TABLE 3.12. Rational Expectations Estimation Results using Env. BC Counts

3.7 Conclusions

In this paper, we studied this last potential channel for beneficial effects of VPRs. In particular, VPRs could induce participant firms to innovate in their environmental technologies, thus lowering their costs of over-compliance. Pollutant reductions generally require costly reformulation of products or production processes, suggesting that over-compliance positively impacts environmental innovation. We examined an unbalanced panel of 127 manufacturing industries defined by 3-digit SIC classifications (SIC codes 200-399) over the period 1989 – 2002.

Our conclusion from this paper is that VPRs may potentially have costs that have not before been recognized or anticipated by scholars or policy-makers. In particular, we find evidence that participation in the 33/50 program may have diverted resources away from longer-term environmental R&D investments, leading to longer-run reductions in our patent count measure of successful environmental research. This outcome requires further study to determine its robustness to different lags of 33/50 program effects on patent outcomes and different models of expectations formation. It also suggests the need for more work to determine any long-run effects of VPRs on ultimate environmental performance.[35] However, if these results withstand further scrutiny and to the extent that environmental R&D is considered the engine of environmental improvement, this paper suggests that VPRs may potentially have an important environmental cost that may or may not outweigh the short-run emission reduction benefits identified in prior work.

[35] Extant studies of 33/50 effects on emissions have focused only on years before the program ended (pre-1996) and hence have not measured any long-run effects, potentially via research channels.

Chapter 4

WHAT MAKES YOU GO BACK HOME? DETERMINANTS OF THE DURATION OF MIGRATION OF MEXICAN IMMIGRANTS IN THE UNITED STATES

4.1 Introduction

The evolution of Mexican migration to the United States is generally understood to be the result of several forces that encourage migration. Theoretical models and recent studies on Mexican migration have suggested a dynamic pattern of cross-border migration in which the economic situation in Mexico and the United States, as well as the presence of relatives in the United States, determine the location and length of stay of Mexican migrants (Massey et al., 1987; Hanson and Spilimbergo, 1999; Lindstrom, 1996).

In this paper, I argue that the relation between duration of migration trips and forces that spur migration (i.e. individual characteristics, location choices, labor market conditions), are important and interdependent. Although the duration of migration trips will depend on the locational choice, the type of employment available, *paisanos*[36] living in the area; where the migrant chooses to go may be dependent on how long they are planning to stay. Therefore, a clear understanding of the determinants of trip duration is crucial for evaluating future U.S. migration policies. In this paper, I present a simple theoretical model of migrant's trip duration that incorporates social and economic factors in Mexico and in the United States.

This paper fills a gap in the migration literature by analyzing the factors that increase

[36] Paisanos refers to individuals from the same origin community in Mexico but are not considered close relatives.

or decrease the length of stay of Mexican immigrants in the United States, an issue widely ignored in the migration literature. A notable exception is Angelucci (2005) where she looks at the probablity of staying of a Mexican Immigrants in the United States using a discrete choice model. Conversely, this study uses a porportional hazard model to estimate the hazard of returning to Mexico. Moreover, this study uses individual level data compared to the aggregated data level used in Angelucci (2005); therefore this study analyses the effect of expected wages and distance traveled on the duration of migration trips that is not feasible when data is at an aggregated level.

By analyzing how much time individuals spend in the United States, we can describe how the duration of migration varies across individuals and in subsequent migrations. This study will examine whether demographic characteristics, economic conditions or social networks drive the duration of Mexican immigrants in the United States. Moreover, two new variables are introduced in the analysis. First, the distance in miles between the origin state in Mexico and the destination state in the United States. This distance accounts for the transportation costs an individual incurs when migrating. Secondly, instead of just using the average U.S. real wage, I use an expected U.S. wage for each migrant, which is a function of unemployment rate level and the average U.S. real wage. This study also examines whether these durations have changed across migration trips and whether the characteristics that drove the duration of the first migration trip are the same for the last migration trip undertaken by the migrant.

The sample is drawn from the Mexican Migration Project (MMP)[37] which is a survey

[37] The survey can be found at: http://www.pop.upenn.edu/mexmig

constructed jointly by research centers based in Mexico and the US from 1982 to 1999. This study focuses on the migration experience and social characteristics of Mexican migrants who have migrated to the United States and some of them have never returned to Mexico. I use is a cross-section of 2375 individuals aged 15-64, who report the duration of their first trip to the United States. A second sample is also a cross-section of 2658 individuals who report the duration of their last trip to the United States.

This paper uses the Cox proportional hazard model to estimate the impact of characteristics of the individual, household, destination and origin areas, and U.S. migration policy on the hazard of returning to Mexico. I find that the U.S. expected wage and changes in the United States migration policy have a significant impact on the hazard of returning to Mexico, especially for the last migration sample.

The use of the Cox proportional hazard model offers several advantages. First, the Cox proportional hazard model makes use of all the available information on observations for Mexican immigrants who migrated to the US and have not returned to Mexico at the time of the survey (8% of migrants in the first migration sample never returned to Mexico and 25% of migrants in the last migration sample never returned to Mexico), as well as for those who have already returned to Mexico. Second, the model yields estimates of baseline hazards, establishing, for example, that the migrants who are proficient in English have a lower hazard of return relative to those do not speak or understand English[38]. Third, the model allows very flexible specifications; it appears that

[38] We cannot tell from the survey whether the English proficiency reported was gained during the migration trip or it was the case that the migrant had gained those skills before making the trip. If it is the case that English proficiency was gained during the migration trip then we may have endogeneity issues between the adquisition of the English language skills and the expected duration of stay.

the effect of migration policy on the hazard of return to Mexico falls as the probability of being apprehended increases, while a tighter U.S. migration policy increases the hazard of returning to Mexico. As a consequence, we have an ambiguous effect of the migration policy on the duration of migration by Mexican immigrants. This result is consistent with Angelucci (2005).

This study is organized as follows. First, I present a literature review of studies that have analyzed the determinants for people to migrate. Second, I present a simple theoretical migration model that illustrates the relationship between wage differentials and optimal migration duration. Thirdly, I analyze empirically the individual and economic characteristics that modify the length of stay in the US. Fourth, I provide estimates that suggest Mexican immigrants are deciding to stay longer in the US (therefore forming a more permanent illegal community in the US), or, they are choosing to stay for shorter periods of time (suggesting a repeated migration pattern which suggests that the migrant moves back and forth between Mexico and the United States). Finally, policy implications and future lines of research are presented in the conclusions of this paper.

4.2 Literature Review

Migration literature has indicated three major reasons for people to migrate to the US. First, the difference in real wages between the US and Mexico is considered the most important reason why Mexicans choose to migrate (Hanson and Spilimbergo, 1999; Chiquiar and Hanson, 2002). Migration occurs under the assumption that observed behavior is preceded by a desire to migrate. This is based on the seminal work by Stjaastad (1962) in which migration is viewed as an investment decision. An individual decides to migrate if the expected discounted difference in the income stream between two places exceeds moving costs.

The second reason is the existence of job and social networks among Mexican immigrants in the US. Previous literature suggests that the most important determinant of immigrants' locational choices within the United States is the presence of earlier immigrants. For example, we would expect that the probability of an immigrant living in a certain city is positively correlated with the fraction of the same ethnic population that resides in the area. (Bauer et al., 2003; Bartel 1989). Moreover, highly educated immigrants tend to be less geographically concentrated than less educated immigrants. Chiswick and Miller (2004) find geographical agglomeration among international migrants by language. On the other hand, Bartel (1989) finds that highly educated immigrants tend to be less geographically concentrated than less educated immigrants. Jaeger (2004) finds that migrants are more likely to locate in areas/neighborhoods with a high proportion of foreign born individuals. Finally, Garcia (2005) found three distinct yet disconnected subnetworks: a traditional subnetwork, a church subnetwork and a contact subnetwork.

Third, the demand for unskilled labor that exists in the United States. Labor market conditions affect the locational choice and at the same time impact the length of stay. Bartel (1989) found that Hispanics are less likely to live in areas with high unemployment rates. High unemployment levels in the United States are likely to negatively impact the amount of time Mexican immigrants stay in the US.

Despite a wide literature on the incentives of Mexican Immigrants to move to the US, we know little about the determinants of duration in the US by Mexican immigrants. Return migration may occur despite a positive wage differential for two reasons. First, we can assume a relatively high preference for consumption at home; this means that the preference to stay in Mexico for longer periods of time is higher when the family resides in Mexico and owning a property in Mexico[39]. Secondly, the existence of a higher purchasing power of the dollar in Mexico. In the past, Mexico has suffered peso collapses that suddenly increased the purchasing power of the dollar. These conditions may incentive an *earlier* return since his savings increased in relative terms. An important issue is the role that expected wages play in the optimal migration duration. Intuition suggests that the optimal duration of the migrants in the host country increases if the expected wage increases.

In the past, temporary migrations were frequent, among Mexican immigrants. This pattern may be changing due to the tightening of the border between Mexico and the United States. Recent evidence suggests that an increasing proportion of migrants eventually settle permanently in the United States (Vernez and Ronfelt, 1991). Moreover, Angelucci (2005) finds that while increased border enforcement discourages migrants from crossing into the United States, it may discourage the return to Mexico of migrants already in the United States. Even so, temporary migrants still constitute a significant portion, if not a majority, of Mexican migrants to the United States. This change in the

---

[39] It is important to stress the fact that owning property in Mexico may be endogenous because the migrant accumulates wealth during his stay and may delay his return to Mexico.

pattern of migration trips to the United States implies that the United States is experiencing an increase in the number of illegal migrants residing within the United State's borders. The existence of an illegal community in the United Sates has made migration reform a top policy issue, which proposals have been made regarding how to deal with this community.

There is previous research that analyzed international return migration intentions among guest workers in Germany. Dustmann (2003) highlights the existence of further motives for a return migration. These are a high purchasing power of the host country currency in the migrant's home economy, and higher returns in the home economy to human capital accumulated during the stay in the host country. Waldorf (1995) shows that return migration is influenced by residential and job satisfaction and stage of the life cycle. In the case for Mexico, Lindstrom (1996) finds that migrants from economically dynamic areas in Mexico with favorable opportunities for employment and higher returns to small capital investment have a larger incentive to come to the US. This is counterintuitive because we would expect that the lack of opportunities in Mexico is the main driving force for migrants to cross the Mexico-U.S. border. But at the same time, we see that, illegal migrants come from different social and economic classes as well as from a diverse array of locations.

4.3 The Conceptual Framework: Migration Duration

The purpose of this section is to construct a simple theoretical model of migration that illustrates the relationship between wage differentials and optimal migration durations.[40] The economy starts at some initial point in time, $t=0$. The individual dies at time T. The migration pattern between Mexico and the US results from individual decisions of potential migrants who are individuals age 15 to 64. Also, assume that migration results from a positive difference in real wages between US and Mexico and the existence of social networks in the US. Individuals choose the optimal migration duration given a positive wage differential between Mexican and US wages $\left(w^{mx}<w^{us}\right)$. Assume wages in Mexico and US are constant throughout the lifetime and that there is a continuum of migrants with different abilities and heterogeneous migration costs. The utility function also depends on the consumption preferences by the migrant $\left(c^{us}, c^{mx}\right)$ and assumes that the migrant will have an appetite for consumption in both places. Every month, a migrant decides whether to stay in the US or return to Mexico. If the immigrant decides to return to Mexico, then temporary migrations occur. It is assumed that the migrant chooses the optimal duration in the United States and will return at $\hat{t} \in (t,T)$. However, given the parameters of the model, different abilities will be associated with varying optimal migration durations, including $\hat{t}=t$ and $\hat{t}=T$, which refer to permanent migrations. Permanent migrations occur when individuals choose to stay in Mexico or in the US for their lifetime.

This study only considers the case of interior solutions, which is where we have

[40] This model is based on the one developed by Dustmann and Kirchkamp (2001). This model is more simplistic since we are not interested in the activity a migrant pursues once thay have returned to their place of origin. However, the model still captures the essential tradeoff between staying longer in the US and returning to Mexico.

temporary migrations. To simplify the analysis there is no uncertainty in the model. To represent the choice problem, assume the existence of a utility function representing individual preferences. The migrant's lifetime utility function is given by

(4.1) $$U = \sum_{\tau=1}^{\hat{t}} u\left(\mu^{us}, c^{us}\right) + \sum_{\tau=\hat{t}+1}^{T} u\left(\mu^{mx}, c^{mx}\right)$$

since it was assumed that wages $\left(w^{mx}, w^{us}\right)$ are constant throughout a lifetime. The utility function simplifies to:

(4.2) $$U = \hat{t}u\left(\mu^{us}, c^{us}\right) + (T-\hat{t})u\left(\mu^{mx}, c^{mx}\right)$$

where $u(\cdot)$ are the utility functions in Mexico and the United States. Denote preference parameters as μ^{mx} and μ^{us} , which indicate if a migrant prefers to live in Mexico rather than in the US, then $\mu^{mx} > \mu^{us}$ .Assume that immigrants have an appetite for consumption in both places, that is $u_c > 0$ and $u_{cc} < 0$. The maximization problem is represented as:

(4.3) $$\max_{c^{us}, c^{mx}, \hat{t} \in (t,T)} U\left(c^{us}, c^{mx}, \hat{t}\right) = \hat{t}u\left(\mu^{us}, c^{us}\right) + (T-\hat{t})u\left(\mu^{mx}, c^{mx}\right)$$

$s.t.$

$$\left(T-\hat{t}\right)w^{mx} + \hat{t}w^{us} - \hat{t}pc^{us} - \left(T-\hat{t}\right)c^{mx} - cc = 0$$

$$w^{mx} < w^{us}$$

In equation (4.3) the parameter p denotes the relative price of consuming in United States relative to Mexico. If $p>1$ then consumption in the United States is more costly than consumption in Mexico.[41] The term (cc) denotes the costs of crossing the border, which include transportation costs, ability to cross the border, non-labor income (e.g. income from property owned in Mexico) and forgone income.

Denote the marginal utility of wealth as (θ). Differentiating the associated Lagrange problem with respect to the optimal time of return $\hat{t}$ yields the condition which determines the optimal migration duration:

$$(4.4) \qquad \theta\left[\left(w^{us}-pc^{us}\right)-\left(w^{mx}-c^{mx}\right)\right]-\left[u\left(\mu^{mx},c^{mx}\right)-u\left(\mu^{us},c^{us}\right)\right]=0$$

The first term of equation (4.4) $\left[\left(w^{us}-pc^{us}\right)-\left(w^{mx}-c^{mx}\right)\right]$ represents the benefit of remaining an additional month in the United States. Assuming that $\left(w^{us}-w^{mx}\right)>\left(pc^{us}-c^{mx}\right)$, $w^{us}>w^{mx}$ and $p>1$ we expect this term to be positive (i.e. each month spent in the United States. increases the migrant's lifetime wealth) but decreases in $\hat{t}$. The second term of equation (4.4) represents the costs of staying one additional month in the United States and is increasing in $\hat{t}$. Unfortunately, this second term cannot be signed because although μ^{mx} is greater than μ^{us} the model makes no assumption about the ordering of c^{mx} and c^{us}. The difference between benefits and costs

[41] This also accounts for the purchasing power of the dollar in Mexico.

of staying one additional month in the United States decreases in $\hat{t}$. The optimal time of return is when the benefits of staying one more month in the United States are equal to the costs of doing so.

Comparative static are derived using equations (3) and (4) and the first order conditions for c^{mx} and c^{us}. The change in the optimal migration duration as a result of the changes in wages in the US is summarized as follows:

(4.5) $$dt = \psi_1 dw^{mx} + \psi_2 dw^{us}$$

where ψ_i combines the partial derivatives of (3) and (4) with respect to θ, $\hat{t}$, w^{mx}, w^{us}. If we assume that there is an increase in the wage differential then we expect an increase in the migration duration. However, the income effect is negative because the value of staying in the US one additional month decreases as total wealth increases, leading to a reduction in the optimal migration duration.[42] Consequently, the theoretical model does not tell us if we should expect an increase or decrease in the duration of migration trips due to an increase in the wage differential. Therefore, I explore the empirical implications of this simple theoretical model. While there are other effects to analyze theoretically, they are left for further research on the topic. For example, the effect on the optimal migration duration of an increase in Mexican real wages, the effect of an increase in border enforcement, the effect of an increase in migration experience.

---

[42] Note that the theoretical model presented above refers only to the duration of a single trip, and not to lifetime participation in the US labor market or to the frequency of trips.

4.4 Empirical Model

This study focuses on understanding the determinants of migrants' trip duration. This topic has been largely ignored in the migration literature, which has mainly focused on describing the individual characteristics of migrants and estimating the number of entering migrants in to the United States. (Bean et al, 1987; Durand and Massey, 1992). A notable exception is Angelucci (2005) where she looks at the probablity of staying of Mexican Immigrants in the United States given an increase in border enforcement. Using a discrete choice model; she found that while increased border enforcement discourages migrants from crossing into the United States, it may discourage the return to Mexico of migrants already in the United States. Similarly, this study uses a porportional hazard model to estimate the hazard of returning to Mexico. However, this study uses individual level data compared to the aggregated data level used in Angelucci (2005); which allows analyzing the effect of expected wages and distance traveled on the duration of migration trips that is not feasible when data is at an aggregated level.

The main hypothesis of this paper is that migrants make migration decisions based on the expected wages they will receive in the destination communities, as well as costs of migration. Both factors are expected to be positively related to the length of migrants' stay in the United States. Ideally, Mexican migrants prefer to spend as little time as possible away from home and yet accumulate enough savings during their stay in the United States. Therefore, when Mexican migrants have saved or remitted money equal to some target, they return to their places of origin in Mexico.

In general, the expected length of time a migrant spends in their locational choice on a given trip increases with higher migration costs because the migrant will minimize costs by having a single longer trip than several short trips. On the other hand, it is expected to decrease with lower destination wages because the migrants' costs will exceed the benefits of remaining in the United States.[43] Finally, the expected length of the trip increases with tighter border controls and illegal status. This effect has the same rationale as minimizing migration costs. Every time a migrant crosses the border, he inccurs explicit costs (e.g. coyote fees) and implicit costs (e.g. days waiting to cross the border, time spent crossing the border, risk his life). If these costs increase, then migrant is likely to stay for longer periods of time.

4.4.1 Migration Characteristics

The data used in this paper comes from the Mexican Migration Project (MMP) constructed jointly by researchers based in Mexico and the Unites States from 1982 to 1999[44]. In the MMP, a number of communities were surveyed each year. Each community was surveyed only once, obtaining a retrospective history of migration patterns. The selected communities are diverse in size and economic base; they encompass small agricultural towns, mid-sized towns and metropolitan areas located primarily in the western part of Mexico, which has been characterized as the major supplier of Mexican immigrants. In each community, representative households are

[43] An increase in the host country wage increases the marginal value of staying in the host country (relative wage effect) but, at the same time, decreases the marginal utility of wealth (income effect). Migrants may return earlier, should the wage level in the host country increase. Therefore, there is an ambiguous effect with a change in destination wages.

[44] The survey can be found at: http://www.pop.upenn.edu/mexmig

selected through simple random sampling.[45]

I focus on the migration experience and social and economic characteristics of individuals aged 15 to 64[46] who have migrated to the United States. The migration period considered in this study from 1963 to 1999 is called the modern period of illegal migration, which refers to migrations that occurred after the *Bracero* Program ended.[47] Moreover, the sample period includes three major changes in the US immigration law: The Immigration Reform and Control Act (IRCA) in 1986, the Immigration Act (IA) of 1990 and the Illegal Immigration and Responsibility Act (IIRA) of 1996. Finally, the most recent Mexican economic collapses are also included in the sample: the peso collapses in 1982 and 1994.[48]

The IRCA contained four main provisions: (1) sanctions were introduced on employers hiring illegal immigrants; (2) the Border Patrol resources were increased; (3) an amnesty was provided for undocumented immigrants who could prove they had resided continuously in the United States from 1982; (4) a special legalization program was implemented for undocumented agricultural workers. The next legislative step was the passage of the 1990 Immigration Act. While its main provision was to introduce a yearly cap on total legal immigration to the US, it is important illegal Mexican migrants because the Act provided increased resources destined to enforcement on the Mexican

[45] In some cases, the entire town was surveyed. In large urban cities, however, this procedure is infeasible therefore only demarcated and sampled specific working-class neighborhoods were included in the sample (Durand and Massey, 1992).
[46] This is the age at which the individual migrated. This age can differ from the one reported in the survey because the age reported at the survey is the current age.
[47] Illegal Migration began to rise after the end of the Bracero Program (1942-1964), which permitted farm laborers from Mexico to work in the US agriculture on a temporary basis. Laborers were required to return to Mexico after completing their contract work (see e.g. Hanson and Spilimbergo, 1999).
[48] For a more complete treatment of US migration legislation and Mexican economic contractions, see Massey et al. (2002), Hanson and Spilimbergo, 1999 and Angelucci, 2005.

Border.[49] Finally, the Illegal Immigration and Responsibility Act, passed in 1996, mainly increased the penalties to those smuggling immigrants to the United States.

Based on the MMP survey, I can identify complete durations of Mexican immigrants in the United States. This allows me to test the hypothesis proposed in the theoretical model that Mexican immigrants reduce their migration duration in response to higher wages in the United States, or increase their migration duration in response to higher migration costs (e.g. increased border enforcement). The sample consists of male and female immigrants, with the majority of immigrants crossing the border illegally.[50] I created two samples, one that refers to the social and economic characteristics during the first migration trip undertaken by individuals, and the other refers to social and economic characteristics during the last migration trip.

[49] During the 1990's, a series of local operations against illegal border crossing were put in place by the Border Patrol. The main feature of these operations was to discourage illegal border crossing.

[50] In the sample, 45% of migrants report crossing the border illegally, that is they don't have the necessary papers to cross the border. The rest cross the border legally but become illegal migrants by staying after the expiration date on their visas.

Duration in Months	Less than 6 mos	6 mos – 12 mos	13 mos – 36 mos	37 mos – 60 mos	60+ mos	Total
FIRST MIGRATION						
Number of Migrants	871	728	465	137	174	2375
Percentage	0.37	0.31	0.20	0.06	0.07	1.00
Mean						26.69
Standard Deviation						57.16
LAST MIGRATION						
Number of Migrants	1091	658	433	158	318	2658
Percentage	0.41	0.25	0.16	0.06	0.12	1.00
Mean						22.91
Standard Deviation						54.21

Source: Mexican Migration Project, Migration File

TABLE 4.1. Return Frequencies Summary Statistcs

Table 4.1 presents summary statistics for durations of the first migration and the last migration of Mexican immigrants. It also presents the percentage of individuals that return to Mexico in the indicated period of time. There is a high percentage of individuals returning within six months of arriving to the United States in the last migration than in the first migration. Moreover, there are a higher percentage of individuals that stay more than five years in the last migration than in the first migration. This accounts for the fact that the last migration sample has 672 right censored observations. I expect that the empirical results are not driven by the behavior of long-term stayers because the sample still portrays both types of migration trips (e.g. return and permanent migration). Furthemore, the number of censored long-term migrants (duration greater than five years) in the sample is relatively small.

4.4.2 Definition of Variables/Covariates

Table 4.2 and 4.3 define the variables used in the estimation and presents summary statistics for the first and last migration samples. Characteristics of the individual include age of individual when migration occurred (***Age***), marital status and place of residence of the spouse (***married MX*** and ***married US***), number of children aged up to 15 years old (***minors***)[51], occupation, and education. These last variables (e.g. occupation and education) are a measure of human capital and are constructed using flexible education[52]

[51] Unfortunately, I can't distinguish whether the children reside in Mexico or in the US. The survey only reports number of children and their ages.
[52] Several specifications for the education dummy were tested. The most flexible was using a dummy for the highest level of completed schooling to the most restrictive using a dummy to indicate that the migrant completed elementary school. The specification used is the most parsimonious and the best that describes the characteristics of the sample.

and occupation specifications. The reason for a flexible specification is to control for the heterogeneity in the ability of migrants to obtain a job in the US. Finally, the migrant reports his English proficiency (***ESL***); that is if they understand, speak, write and read in English. They have a score from 1 to 4, where 4 is the highest score and they must read, write, and speak English and 1 is the lowest score. *ESL* is a dummy variable that takes the value of one if they report a score of 4 and zero otherwise.

Migration costs are defined as whether the city of origin is considered an urban (***Urban***) or a rural area.[53] This distinction of migration costs is due to the fact that it is more likely that the migrant in an urban area has more options to travel than a migrant from a rural area. The apprehension rate for the year they crossed the border (***Apprehension***)[54] and the distance in miles between origin state capital in Mexico and destination cities in the United States (***Distance***)[55].

---

[53] According to the Mexican Census, an area is considered urban if its population is greater or equal to 50,000.
[54] This average is the ratio between total apprehensions and the total undocumented migrations.
[55] If the destination city in the United States is not reported then we use the distance traveled between origin state capital in Mexico and the destination state capital in the United States.

Variable Characteristic	Variable Name	Definition	First Migration Sample Mean	First Migration Sample S.D.	Last Migration Sample Mean	Last Migration Sample S.D.
Individual	Age	Age at time of migration	26	9.252	26	9.010
	Married MX	1 if married and spouse resides in Mexico	0.888	0.35	0.972	0.165
	Married US	1 if married and spouse resides in US	0.089	0.285	0.093	0.291
	Children	Number of children less than 15 yrs old	1.114	0.815	1.133	0.832
Occupation	Agricultural	1 if in Agriculture	0.285	0.451	0.305	0.461
	Professional	1 if practices profession	0.052	0.222	0.056	0.231
	Manufacturing	1 if in manufacturing	0.176	0.381	0.219	0.414
	Unskilled	1 if unskilled laborer	0.163	0.369	0.152	0.359
	Self Employed	1 if self employed	0.206	0.405	0.197	0.398
Household	Mother US	1 if Mother resides in US	0.103	0.304	0.098	0.297
	Father US	1 if Father resides in US	0.324	0.468	0.322	0.467
	Property MX	1 if owns property in Mexico	0.724	0.447	0.732	0.443
Origin	Urban	1 if comes from urban area	0.175	0.380	0.166	0.372
Schooling	Elementary	1 if completed 5th – 6th Grade	0.312	0.463	0.291	0.454
	Some Middle Educ	1 if completed 7th – 8th Grade	0.052	0.222	0.047	0.212
	Middle Educ	1 if completed 9th Grade	0.117	0.322	0.107	0.309
	Some High School	1 if completed 10th – 11th Grade	0.031	0.173	0.028	0.165
	High School	1 if completed 12th Grade	0.045	0.207	0.041	0.198
	Some College Educ	1 if completed 13th – 15th Grade	0.027	0.163	0.025	0.156
	College Educ	1 if completed 16th – 17th Grade	0.024	0.154	0.022	0.146
	Some Grad Educ	1 if completed 18th + Grade	0.009	0.094	0.008	0.091

Source: Mexican Migration Survey

* Series deflated by the US Consumer Index (CPI)

♦ Normalized by 1,000 miles

TABLE 4.2. Summary Statistcs

Destination communities are described by the presence of kin in the United States (*Mother US* and *Father US*), social networks (***paisanos***) and the wage the migrant expects to receive (***expw***). Mexican migrants' response to the higher returns associated with increased trip duration should be influenced by the availability of employment in the destination community[56]. Therefore an expected wage measure was constructed instead of using just the average of the real wage. The expected wage in the United States at the year the migrant decides to migrate, is constructed as $(1-u_{it})(w_{it}/p_{it})$, where u_{it} is the unemployment rate in the destination state i in the United States at time t, w_{it} is the mean wage in the destination state i in the United States at time t, and p_{it} is the US CPI.[57]

United States migration policies are described by ***irca***, which indicates whether the migrant was legalized during the amnesty of 1986, also whether the migration duration occurred after 1986 (***Year 1986***) and 1990 (***Year 1990***) Migrations Acts. Finally, long term savings are described by whether the migrant owns property in Mexico (***property mx***), the total amount of savings brought to Mexico (***saving1*** and ***saving2***)[58], and whether remittances were sent to Mexico during their stay in the US (***Remittances***). I also control for the year migration occurred (*year migration*)

---

[56] In this case there is a disconnect between the theoretical and empirical models since our empirical work does allow real wages to change unlike the assumption of the theoretical model.

[57] Ideally we would like to construct a more individual specific expected real wage rate. It may be possible to predict the potential unemployment rate and wage for a given migrant in a given community. This will potentially reduce the bias produced when using aggregated unemployment rates and wage rates.

[58] Saving1 one denotes those individuals that saved from 500 to 2500 dollars during their trip and saving2 refers to those individuals that saved more than 2500 dollars during their trip. There is potential endogeneity when using these measures because we can expect that savings in the United States may be jointly determined with the duration of stay.

Finally, the dummy (***property mx***) captures the effect of investment in fixed capital assets that can be sold in the future with some gain and proxies for long term savings of migrants. This is in addition to the amount brought to Mexico after the trip has ended. Agricultural land and residential real state are two of the most common forms of fixed capital in which migrants invest, as shown in Table 4.4. Therefore, *property mx* equals one whenever the migrant reports ownership of agricultural land and/or residential real state.

Variable Characteristic	Variable Name	Definition	First Migration Sample Mean	First Migration Sample S.D.	Last Migration Sample Mean	Last Migration Sample S.D.
Destination	Paisanos	1 if members same community reside same destination	0.634	0.482	0.637	0.481
	Exp Wage *	Expected Wage in destination	9.858	1.211	9.531	1.361
Crossing Border Costs	Distance ♦	Distance in miles from origin to destination community	1.572	0.484	1.611	0.475
	Apprehension Rate	Probability of Apprehension	0.315	0.048	0.288	0.057
Migration Policy	Year Migration	Year they migrated to US	1981	7.756	1986	7.942
	IRCA	1 if legalized by IRCA 1986	0.928	0.259	0.158	0.364
	Year 1986	1 if crossed border after 1986	0.174	0.379	0.234	0.423
	Year 1990	1 if crossed border after 1990	0.129	0.335	0.364	0.481
Savings	Savings 1 *	1 if saved from 500 to 2500 usd during trip	0.261	0.440	0.258	0.437
	Savings 2 *	1 if saved more than 2500 usd during trip	0.186	0.389	0.187	0.390
	Remittances	1 if sent remittances to Mexico while in US	0.661	0.473	0.657	0.475

Source: Mexican Migration Survey

* Series deflated by the US Consumer Index (CPI)

♦ Normalized by 1,000 miles

TABLE 4.3. Summary Statistcs

Spending Category (%)	First Mgration Sample	Last Migration Sample
Productive Capital	27.83	27.62
Vehicle	0.17	0.1
Consumer Goods	3.58	3.57
House/Lot	11.37	11.96
Home Construction/Improvement	5.52	5.38
Family Mantainance	45.09	44.89
Recreation	0.16	0.16
Debt	3.41	3.54
Savings	1.09	1.02
Other	1.78	1.76

Source: Mexican Migration Project, Housefile

* Bussiness, Tools, Farmland, Livestock

TABLE 4.4. Use of Remittance Income

The presence of kin (*mother US* and *father US*) and persons from the same community (*paisanos*) in the United States measure the extensiveness of social ties of Mexican immigrants. Together with marital status and community based migration networks increase individuals migrants' flexibility and freedom in the choice of migration strategies because both sets of factors facilitate long-term migration to the US as well as frequent cross border movement. The human capital and migration experience variables are set to be equal to the values they assume at the start of the trip.

Migration duration is a function of the migration policy and migration costs at the time of the trip. Therefore, in order to measure the impact of changes in the duration by changes in the US migration policy and border enforcement, I control for legalization sponsor (*irca*), year migration occurred, whether they migrated after Migration Acts were in place and the rate of apprehensions the year they migrated. It is expected that an increase in US Border enforcement significantly reduces the flow of Mexican immigrants since migration costs are higher. Hence, migration durations are longer when the immigrant faces higher migration costs. Consequently, the individual expectations of future migration costs also affect the optimal duration of subsequent migrations.

Both samples (first and last migration) are fairly similar when considering individual characteristics (age, education, occupation, etc.) It is interesting to note that both samples are consistent with common characteristics of Mexican immigrants described in the literature.[59] They are poorly educated, younger and more likely to be males.[60] However,

---

[59] See Borjas (1991) and Cuecuecha (2003) where they highlight a decline in immigrant skills, prominently males and young.

[60] 95% of the individuals in the sample are males.

when we examine the nature of immigrant flows from Mexico to the United States, other important features are highlighted in the sample. First, past migrants were highly concentrated in only two states (California and Texas) and accounted for 81% of Mexican migrations. In recent migrations there is more variation across states. Migrants mainly choose to locate in California, Texas, Illinois, Florida, and Arizona. This partly explains the increase in the mean distance measure from the first migration to the last migration, together with the fact that more recent migration comes from southern states in Mexico (e.g. Oaxaca, Puebla) that are farther away from the border.

Another important difference is that 92% of individuals in the first migration sample reported to be legalized through IRCA. This legalization process may have not occurred during the first migration. Also, there is a higher percentage of migrants that crossed the border after 1986 in the last migration than the first migration. This is somewhat expected because the last migration is definitely more recent than the first one. For example, 23% of Mexican immigrants in the last migration trip sample made the trip after 1986 compared to 17% of Mexican immigrants in the first migration trip sample.

4.4.3 The Statistical Model

My choice of model for the analysis of returning migration is the Cox proportional hazard model to assess the impact of characteristics of the individual, the origin communities and the destination communities on the hazard of returning to their origin communities. The choice of a hazard model comes naturally because this framework requires one of two possible outcomes for each individual in the sample. The first one is that the individual was in the US when surveyed, which means that the duration had not ended. This outcome in the hazard framework is considered as right censored. The other outcome is that the individual returns to Mexico, in this case it is considered as a "failure".

The instantaneous hazard rate of return migration at time t, conditional on survival to time t can be written as

$$h(t;x_i) = h_0(t) * \exp(x_i'\beta) \quad (4.6)$$

where $h(t;x_i)$ is the hazard of return migration at time t for a migrant described by a vector of coefficients β associated with covariates that characterize the social and economical characteristics of migrants in the sample, where $h_0(t)$ is considered the baseline hazard rate. The crucial assumption in the Cox proportional hazard model is that the effect of the covariates is proportional over the entire baseline. Since the baseline hazard gives the shape of the hazard function, under the Cox proportional model it will be the same for any given individual. Therefore, $h_0(t)$ is the same for all individuals and only the level of the hazard function $\exp(x_i'\beta)$ is allowed to differ across individuals. While this is one of the simplest duration models available, and can be considered as a baseline for future estimations with other duration models, it is sufficiently rich to capture many data properties.

The Cox proportional hazard model has several features that make it an attractive statistical framework for the problem at hand.[61] The most obvious is that it incorporates both the social components that affect the return migration decision and the economic characteristics that modify the length of stay. Secondly, the Cox proportional hazard model exploits all the available information in observations that are right censored.

The risk of return migration is allowed to vary over time and with variation in the covariates. Hazard ratios (exponentiated coefficients) greater than one are indicative of increasing hazard rates and thus are associated with a reduction in the expected time in the US until returning to Mexico. Hazard ratios less than one imply that migrants postpone their return to Mexico, consequently having longer trips.

Another desirable feature of the Cox proportional hazard model is that it readily yields an estimate of the underlying baseline hazard function. The estimation of the baseline hazard function enables us to identify the average length of stay of Mexican migrants in the US. This information is useful for policy makers who would like to avoid a more permanent illegal migration in the United States.

Although the model is dynamic, the data are recorded in discrete intervals, particularly in months. As a result, there are numerous migration spells of the same duration. Duration times are handled using the Peto-Breslow approximation procedure.[62] This approximation takes into account all the individuals that exit at the same time and adjusts the likelihood function. This implies that the likelihood function can be approximated as

---

[61] The proportional hazard model is a common choice for modeling durations because it is a reasonable compromise between the Kaplan-Meier estimator (see below) and the possibly excessively structured parametric models (Greene, 2003).

[62] Described in Kalbfleish and Prentice (1980).

(4.7) $$L = \prod \frac{\exp\left(\sum_{j \in Di} x_{ij}^{'} \beta\right)}{\left[\sum_{l \in R_i} \exp\left(X_l \beta\right)\right]^{m_i}}$$

where i is an ordered failure times $t(i)$, $(i = 1,...,k)$, D_i is the set of observations j that fail at time $t(i)$, m_i is the number of individuals who exit at time $t(i)$, and R_i is the set of all observations l that are at risk to exit at time $t(i)$.

Unfortunately, this specification of the partial log-likelihood function does not explicitly account for the potential effect of unobserved heterogeneity on the hazard rate, which is a limitation of the present approach. The problem of heterogeneity in duration models can be viewed essentially as the result of an incomplete specification. Individual specific covariates are intended to incorporate observation specific effects. With this framework, the best way to account for individual heterogeneity is to include a diverse array of individual covariates in the hazard model which control for individual characteristics as well as household characteristics. Meyer (1990) suggests that explicitly modeling unobserved heterogeneity has little effect on the estimated coefficients in a model in which the baseline hazard rate is allowed to be non-parametric. Nevertheless, I intend to explore the use of duration models that account explicitly for unobserved heterogeneity in future research.

One of the models that can be used to account for unobserved heterogeneity is the mixed proportional hazard model (MPH) which was proposed by Lancaster (1979). The MPH model controls for the unobserved heterogeneity among Mexican migrants and also for immigrants who have made several trips to the United States illegally. Within this framework, I will study the impact of the length of previous trip duration on the length of the next migration duration trip. This is to address the issue that the estimate of the effect of the previous duration is biased if one ignores the spurious dependence from related unobserved determinants. The MPH model identifies a "causal" effect of a realized past trip on the current trip by including the realized past trip as an additional covariate in the hazard for the current trip.

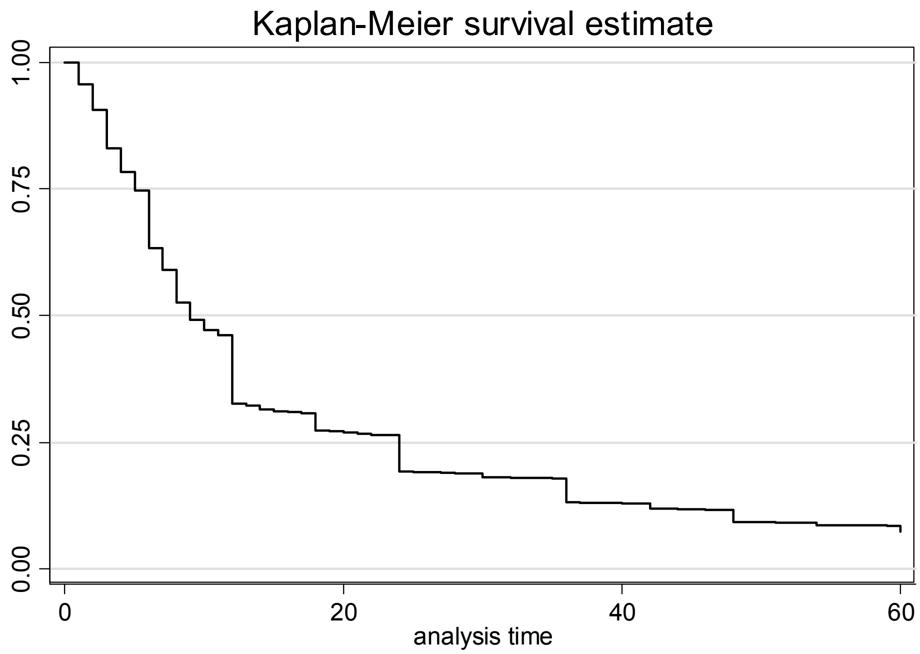
4.4.4 Diagnostics and Specification Analysis

In this section, I undertake a series of diagnostic and specification tests of the duration data to provide a context for the estimation of the hazard rate. The purpose of this graphical analysis of the data is to distinguish the best functional forms and the homogeneity of the observations. I use the Kaplan-Meier estimator (also called the product limit estimator), which is the empirical survival function:

$$\hat{S}(t) = \pi(n_i - h_i)/n_i = \pi(1-\hat{\lambda}_i) \tag{4.8}$$

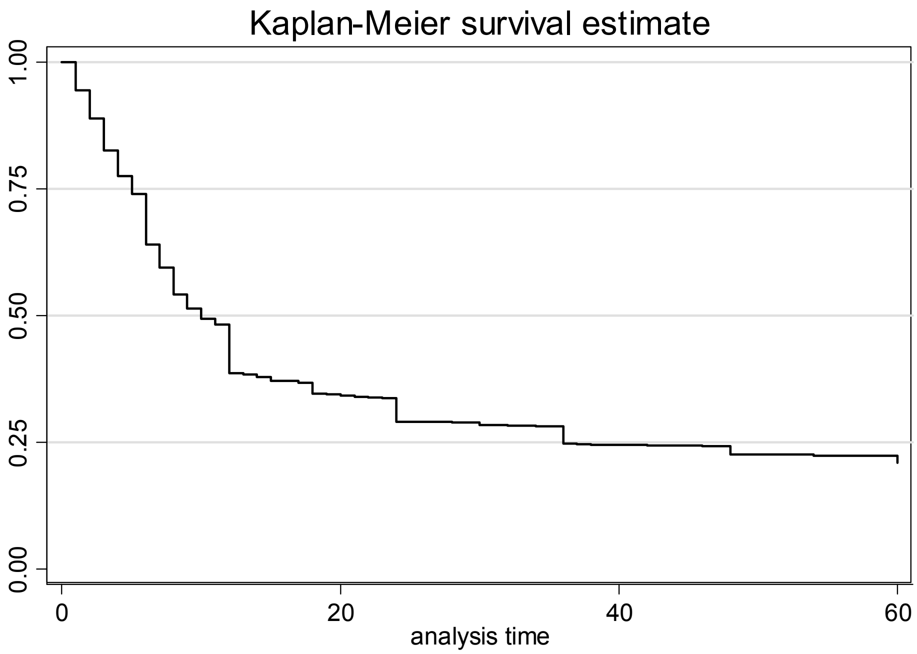
where n_i is just the number "at risk" just prior to time t_i and h_i, the number of failures at time t_i. Therefore $\hat{\lambda}_i$ is the number of "failures" at duration t_i divided by the number "at risk" at duration t_i. I define failures as those migrants that returned to their origin community.

It is immediately apparent, as shown in Figures 4.1 and 4.2, that there is a negative duration dependence, which means that the probability that the duration of the trip ends shortly increases as the trip increases in one more month of stay. Comparing, Figure 4.1 and 4.2, we see that there is a higher probability of returning to Mexico for those individuals migrating the first time as the migration duration increases. Furthermore, it is clear that we do have right censored observations in the last migration sample because the estimates never go all the way down to zero as the first migration estimates do. The graphs also show that the most common returning point occurs at the beginning of the trip. This highlights the temporary migration pattern of Mexican immigrants.



The horizontal axis displays the number of months of trips for the first and last migrations respectively.
The vertical axis displays the Kaplan-Meier survival estimates.

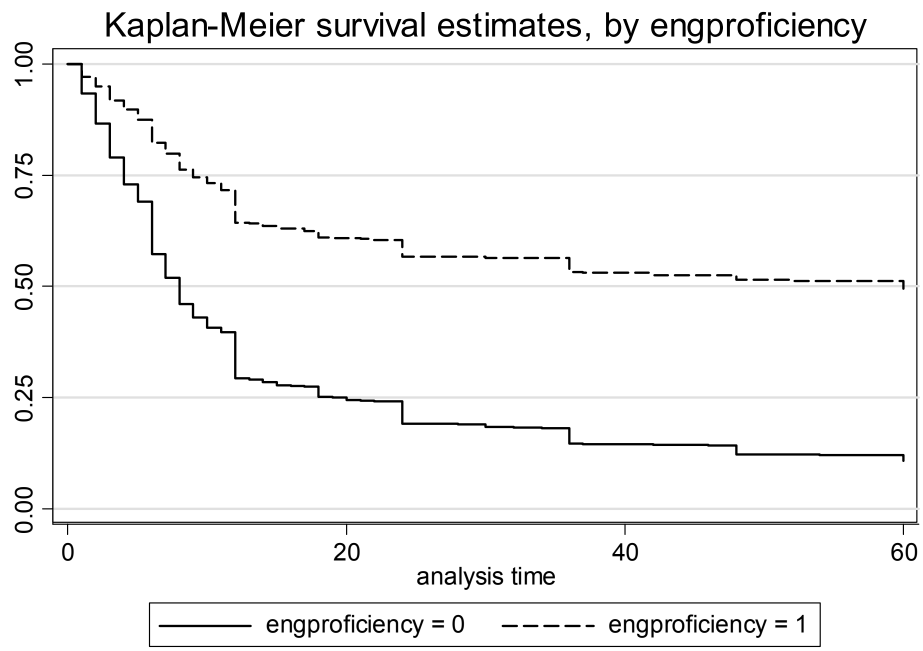
FIGURE 4.1. Kaplan –Meier survival estimate for first migration sample



The horizontal axis displays the number of months of trips for the first and last migrations respectively. The vertical axis displays the Kaplan-Meier survival estimates.

FIGURE 4.2. Kaplan –Meier survival estimate for last migration sample

From this nonparametric analysis, we can test whether the effect of the covariates is proportional over the entire baseline. When the sample is divided by English proficiency (*ESL*) we find that there are different survival functions for each sub-sample (also called strata). In Figure 4.3, we see that those migrants that report being proficient in English have a lower hazard of returning to Mexico. Therefore, we need to estimate the model for each migration trip and stratify the sample by English proficiency.



The horizontal axis displays the number of months of trips for the first and last migrations respectively.
The vertical axis displays the Kaplan-Meier survival estimates.

FIGURE 4.3. Kaplan –Meier survival estimate for last migration sample.
Strata ESL

4.5 Estimation Results

The estimates of the determinants of the hazard of returning to Mexico for the first and last migration are summarized in Table 4.5 and 4.6, respectively. Column one refers to the hazard ratio of each variable on the duration of each trip (exponentiated coefficient). Having a hazard ratio statistically significant means that it is significantly different from one, which in turn means that it does affect the length of the trip.

The (exponentiated) coefficient for the indicator variable *married MX* in the first migration indicates that individuals with their spouse in Mexico are 13% more likely to return during the first migration and 58% more likely to return during the last migration. On the other hand, the variable *married US* is not significantly different from one in both trips, meaning that the variable does not affect the length of the trip. The variable *minors* is significantly different from one, where those individuals with children under age 15 are 10% more likely to return to Mexico in both trips. These results are consistent with Massey et al. (1987) and Waldorf (2003) who found that migrants early in the stage of family formation tend to remain in the host country for shorter periods than do married migrants or migrants without children, who are typically younger.

Education shortens the duration of migration trip. For example, those individuals with at least some college education stay for shorter periods of time than those individuals who are less educated. On the last migration, individuals employed in professional occupations are 30% more likely to return and manufacturing workers have shorter trips as well (21% more likely to return). Conversely, unskilled migrant workers (*unskilled*) have longer trips (for the first migration unskilled workers are less likely to return by

27% and for the last trip they are 14% less likely to return) than other migrants in other occupations. These results might imply that for Mexican immigrants the returns to Mexican schooling are higher in Mexico than in the US. This is consistent with Borjas (1987), who found that the US is a magnet for workers with relatively low earning capacities, and attracts workers with below average skills. A possible cause for these results is that educational requirements are lower for the low skilled work they perform in the US than in Mexico. Given a preference for remaining in Mexico, well educated migrants have greater incentive to spend more time in Mexican labor markets and less time working in the US than do less educated migrants. Finally, agriculture workers have a very high hazard of returning on the last migration (58% more likely to return) but on the first migration they are only 6% more likely to return. A hazard greater than one is consistent with the temporal nature of agriculture.

Variable/Covariate	Hazard Ratio	S.E.	z-stat	P-value
FIRST MIGRATION SAMPLE				
Age	1.0063	0.0816	1.79	0.073
Married MX	1.1371	0.1828	-0.49	0.621
Married US	0.9049	0.0317	3.46	0.001
Children	1.1042	0.0798	0.88	0.377
Agricultural	1.0683	0.0975	-1.44	0.149
Professional	0.8470	0.0733	-1.02	0.308
Manufacturing	0.9222	0.0712	-1.77	0.076
Unskilled	0.8640	0.0670	-1.8	0.072
Self Employed	0.8706	0.1795	-0.26	0.793
Mother US	0.9516	0.0564	3.31	0.001
Father US	1.1726	0.0592	3.17	0.002
Property MX	1.1735	0.0472	-4.23	0.000
Urban	0.7721	0.0480	-2.06	0.039
Elementary	0.8954	0.0962	-0.53	0.599
Some Middle Educ	0.9481	0.0672	-1.79	0.074
Middle Educ	0.8713	0.1167	-0.77	0.443
Some High School	0.9059	0.1021	-0.76	0.450
High School	0.9195	0.1620	1.31	0.189
Some College Educ	1.1950	0.2455	3.34	0.001
College Educ	1.6458	0.4245	2.42	0.016
Some Grad Educ	1.7796	0.0582	5.36	0.000
Paisanos	1.2770	0.0171	-1.61	0.107
Exp Wage *	0.9721	0.0404	-1.94	0.052
Distance ♦	0.9180	0.0076	1.56	0.119
Apprehension Rate	1.0118	0.0048	-3.89	0.000
Year Migration	0.9812	0.0943	0.99	0.320
IRCA	1.0899	0.1261	2.64	0.008
Year 1986	1.2931	0.2159	4.01	0.000
Year 1990	1.6760	0.0561	1.64	0.100
Savings 1 *	1.0884	0.0590	0.6	0.547
Savings 2 *	1.0349	0.0436	-1.9	0.057
Remittances	0.9133	0.0816	1.79	0.073
Log-Likelihood				-14747
Number of Observations				2375
Number of Failures				2375

TABLE 4.5. Estimates of the determinants of the hazard of returning to Mexico for the first migration sample

With respect to social aspects of the destination community, only the indicator variables *father US* and *paisanos* significantly differ from one in both samples. The presence of the father in the US increases the probability of return by 17% in the first migration and by 15% in the last migration. The presence of people from the same origin community in the destination area increases the probability of returning by 27% in the first trip and only 16% in the last trip. The prevalence of recurrent migration among people in a community is an indicator of the reach of migration networks, which are instrumental in reducing the costs of migration. A way to proxy for transportation costs (which also count as migration costs) is to use the distance in miles between the origin state in Mexico and the destination state in the US. As expected, the distance decreases the probability of returning to Mexico by 9% as the distance increases 1,000 miles in the first migration and by 14% in the last migration. These findings support the idea that the expected length of stay decreases with lower migration costs.

Next I examine the effects of origin and destination characteristics. Migrants coming from urban areas tend to have longer trips than migrants coming from rural areas. During the first trip, migrants from an urban area are 23% less likely to return and 27% less likely to return in the last migration than those individuals from a rural area. Migrants who own a house/lot or farmland, have a higher hazard of returning to Mexico. A migrant who owns a property in Mexico increases the probability of return by 17% during the first trip and 41% during the last trip. Migrants from rural areas have little incentive to stay in the United States longer than is necessary to meet current income needs, and it is plausible why migrants from urban areas stay longer periods of time in order to accumulate savings.

The expected wage has the anticipated effect, where an increase of 1 dollar in the expected wage decreases the probability of returning, consequently increasing the optimal time of return. Migrants are highly sensitive to occasional increases in the expected wage. The estimated coefficient for the first migration indicates that an increase of the expected wage decreases the probability of returning by 3% and 4% for the last migration.

The estimated hazard ratio for savings shows that accumulation of savings is only significant in the last migration, where the probability of return of those that were able to save while in the US is higher than those that reported no savings at all. The probability of return for those that reportedly saved between 500 and 2,500 dollars is 32% while the probability of return for those that saved more than 2,500 dollars is 22%. On the other hand, remittances are highly significant for the first migration. If the individual sent remittances to Mexico while in the US, the probability of return decreases by 9%. It seems that the length of the last migration is not affected by whether or not the migrant sent remittances to Mexico while in the United States. These results suggest different intentions for each trip. It is feasible that the savings from the first trip are used to cover current basic needs of the household while the savings from the last trip are intended for long-term savings.

Finally, I look at the migration policy variables. During the first migration, the probability of return increases by 29% and 67% whether migration occurred after 1986 and 1990, respectively. The probability of apprehension seems to be insignificant. This shows a repeated migration pattern even though border enforcement was increased. This is somewhat unexpected because in the theoretical model we concluded that the length of migration trips increases when border enforcement is increased, since migration costs increase. The empirical results suggest the contrary, that border enforcement does not affect the return migration pattern for the first migration. On the other hand, an increase of the border enforcement has the opposite effect on the duration of the last migration. An increase of the border enforcement induces longer migration trips. For example, more experienced migrants show a 2% lower hazard of returning to Mexico when the probability of apprehension increases by 1%. Therefore, we can expect a more permanent illegal community of Mexican immigrants as enforcement increases for more experienced migrants.[63] However, I find that immigrants that crossed the border after 1990 in the last migration, the period which border enforcement was increased, are 25% more likely to return to Mexico. Finally, those migrants that were legalized by IRCA 1986 have a lower hazard of returning to Mexico in the last migration but it does not make a difference for the first migration.

[63] This is consistent with Angelucci, 2005.

Variable/Covariate	Hazard Ratio	S.E.	z-stat	P-value
		LAST MIGRATION SAMPLE		
Age	1.0070	0.0027	2.59	0.010
Married MX	1.5841	0.2812	2.59	0.010
Married US	1.2391	0.6707	0.4	0.692
Children	1.1022	0.0321	3.34	0.001
Agricultural	1.5839	0.1516	4.81	0.000
Professional	1.3071	0.1760	1.99	0.047
Manufacturing	1.2199	0.1246	1.95	0.052
Unskilled	0.7352	0.0852	-2.66	0.008
Self Employed	1.1289	0.1166	1.17	0.240
Mother US	0.5902	0.3153	-0.99	0.324
Father US	1.1550	0.0604	2.75	0.006
Property MX	1.4177	0.0846	5.85	0.000
Urban	0.7355	0.0523	-4.32	0.000
Elementary	0.8637	0.0493	-2.57	0.010
Some Middle Educ	0.9994	0.1216	-0.01	0.996
Middle Educ	0.7489	0.0660	-3.28	0.001
Some High School	0.6868	0.1231	-2.1	0.036
High School	0.8628	0.1159	-1.1	0.272
Some College Educ	0.9303	0.1534	-0.44	0.661
College Educ	1.6762	0.2860	3.03	0.002
Some Grad Educ	1.6692	0.4404	1.94	0.052
Paisanos	1.1676	0.0588	3.08	0.002
Exp Wage *	0.9622	0.0155	-2.39	0.017
Distance ♦	0.8619	0.0410	-3.13	0.002
Apprehension Rate	0.9853	0.0067	-2.18	0.030
Year Migration	0.9908	0.0059	-1.53	0.125
IRCA	0.7414	0.0561	-3.95	0.000
Year 1986	0.9802	0.1081	-0.18	0.856
Year 1990	1.2581	0.1695	1.7	0.088
Savings 1 *	1.3298	0.0723	5.24	0.000
Savings 2 *	1.2261	0.0739	3.38	0.001
Remittances	1.0049	0.0527	0.09	0.925
Log-Likelihood				-13058
Number of Observations				2658
Number of Failures				1986

TABLE 4.5. Estimates of the determinants of the hazard of returning to Mexico for the last migration sample

The estimation of the Cox proportional hazard model stratified by English proficiency yields estimates of the underlying baseline hazard and survival function for a typical migrant in each stratum.[64] Figure 4.4 shows the baseline hazard estimates for each stratum. The hazard rate rise rapidly in the first months of the trip and then flatten out to a two very different levels of risk. The lower baseline hazard applies to those migrants who are proficient in English; the higher hazard rate applies to those who do not speak nor understand English. Therefore, those migrants proficient in English have on average longer durations. This may imply that assimilation in the US labor market of Mexican immigrants is easier if they speak English because the communication costs are lower.

[64]Each stratum is composed by migrants proficient in English (speak and understand English) and those migrants who are not proficient. In my estimation last migration sample, 26% of migrants report themselves as proficient in English while in the first migration sample 20% are proficient in English.

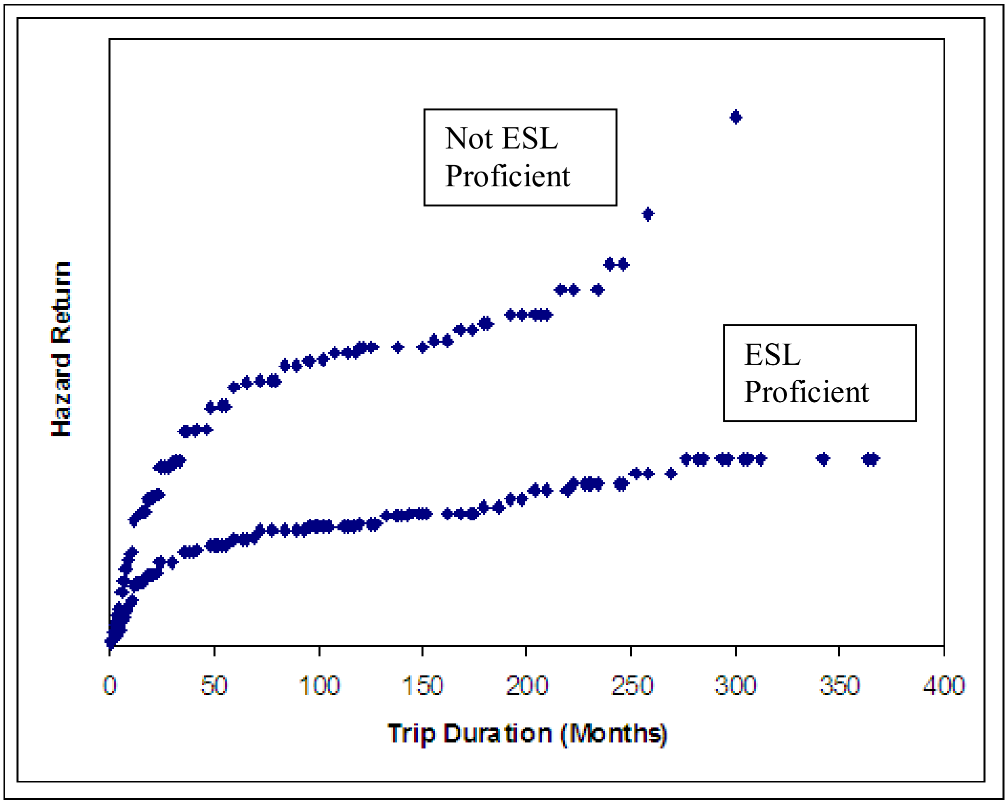


FIGURE 4.4. Estimated baseline hazard for last migration sample

4.6 Conclusions and Avenues for Future Research

This paper developed a simple model, with useful insights, regarding the migration duration of Mexican immigrants. Once in a destination area, temporary Mexican immigrants decide how long they will stay. In making this decision, this paper showed the tradeoff the migrant faces. They weigh the economic benefits of remaining longer against the social cost of living abroad. The analysis shows that an increase in the benefits of remaining in the United States is positively correlated with an increase in the optimal duration migration and the costs.

The empirical evidence presented is consistent with some of the predictions of the simple theoretical model. The exception was with regard to border enforcement, where it was found that an increase in border enforcement does not yield longer migration trips, there is still a temporary migration pattern. This is only the case for the first migration. In the last migration there is indeed a change in the return migration pattern since we found that increased border enforcement increases the duration of the trip. Therefore, we find an ambiguous effect of migration policies on the migration duration.

Empirical results also suggest that Mexican immigrants may in fact increase the length of the trip as a result of an increase in US expected wages. It is misleading to only consider wages as a sole indicator for migration patterns, since it neglects other social and economic factors that are important determinants of the migration duration, such as, social networks in the destination area, family ties in Mexico and communication costs. The important social dimension of migration is confirmed by the significant effect of kinship ties to experienced migrants on the hazard to return during last migration. Mexican migration is both an economic and social process. Once migrants are joined by their spouse and children they gradually develop social and economic ties in destination areas; these ties reduce the likelihood of return.

The savings incentive associated with increased last trip duration is strongest for migrants who can convert current foreign earnings into a source of long term income in their place of origin. On the other hand, savings in the form of remittances during the first trip are intended to cover basic needs of the household. Therefore, employment opportunities in community of origin indicate the degree of likelihood that migration is motivated by the need to cover current household expenses as opposed to the simple desire to accumulate savings.

Future research will address the issue of potential effect of unobserved heterogeneity on the hazard rate. Failure to account for heterogeneity may bias the empirical estimates. Also, whether immigrants stay longer when they have access to social assistance. Borjas (1999) finds welfare-receiving immigrants tend to be clustered in certain locations. Moreover, we can tell whether the magnetic effects of welfare (if they affect duration) differ across migrations under different admission categories.

Chapter 5

DISSERTATION CONLCUSION

In this Chapter general conclusions and implications of this research presented. First, an overview of the research is presented. Second, recommendations for future research are discussed.

5.1 Research Discoveries

In chapter 2, we present empirical evidence of bi-directional linkages between environmental standards and environmental performance, on the one hand, and environmental innovation, on the other. Pollutant emissions and environmental R&D are jointly determined as successful R&D prompts policy change and attendant pollution reductions, and as anticipated policy change (and attendant tightening of pollution standards) spurs new R&D. Specifically, we examine 127 manufacturing industries over the fourteen-year period 1989 – 2002, accounting for the joint determination of research and pollution outcomes.

Our empirical results reveal a negative and significant relationship between emissions and environmental patents, in both directions. Thus, environmental R&D both spurs the tightening of government environmental standards and is spurred by the anticipation of such tightening. Empirical results also suggest that a linear feedback model is appropriate in order to capture the dynamic nature of the links between environmental policy and environmental innovation.

These results suggest that there is a salutary process by which the promise of tightened standards stimulates environmental research, and environmental research, by lowering costs of abatement, stimulates tighter standards. However, the ultimate benefits of tightened pollution standards, due to the resulting stimulus to environmental innovation, appear to be modest. While environmental innovation is found to be a very important driver of long-run pollution reduction, environmental policy plays a role in stimulating environmental research that is statistically significant and not inconsequential, but proportionately not very large.

In chapter 3, we studied this last potential channel for beneficial effects of VPRs. In particular, VPRs could induce participant firms to innovate in their environmental technologies, thus lowering their costs of over-compliance. Pollutant reductions generally require costly reformulation of products or production processes, suggesting that over-compliance positively impacts environmental innovation. We examined an unbalanced panel of 127 manufacturing industries defined by 3-digit SIC classifications (SIC codes 200-399) over the period 1989 – 2002.

Our conclusion from chapter 3 is that VPRs may potentially have costs that have not before been recognized or anticipated by scholars or policy-makers. In particular, we find evidence that participation in the 33/50 program may have diverted resources away from longer-term environmental R&D investments, leading to longer-run reductions in our patent count measure of successful environmental research. This outcome requires further study to determine its robustness to different lags of 33/50 program effects on patent outcomes and different models of expectations formation. It also suggests the need for more work to determine any long-run effects of VPRs on ultimate environmental performance.[65] However, if these results withstand further scrutiny and to the extent that environmental R&D is considered the engine of environmental improvement, this paper suggests that VPRs may potentially have an important environmental cost that may or may not outweigh the short-run emission reduction benefits identified in prior work.

Chapter 4 developed a simple model, with useful insights, regarding the migration duration of Mexican immigrants. Once in a destination area, temporary Mexican immigrants decide how long they will stay. In making this decision, this paper showed the tradeoff the migrant faces. They weigh the economic benefits of remaining longer against the social cost of living abroad. The analysis shows that an increase in the benefits of remaining in the United States is positively correlated with an increase in the optimal duration migration and the costs.

The empirical evidence presented is consistent with some of the predictions of the simple theoretical model. The exception was with regard to border enforcement. Here it was found that an increase in border enforcement does not yield longer migration trips, there is still a temporary migration pattern. This is only the case for the first migration. In the last migration there is indeed a change in the return migration pattern since we found that increased border enforcement increases the duration of the trip. Therefore, we find an ambiguous effect of migration policies on the migration duration.

Empirical results also suggest that Mexican immigrants may in fact increase the length of the trip as a result of an increase in US expected wages. It is misleading to only consider wages as a sole indicator for migration patterns, since it neglects other

[65] Extant studies of 33/50 effects on emissions have focused only on years before the program ended (pre-1996) and hence have not measured any long-run effects, potentially via research channels.

social and economic factors that are important determinants of the migration duration, such as, social networks in the destination area, family ties in Mexico and communication costs. The important social dimension of migration is confirmed by the significant effect of kinship ties to experienced migrants on the hazard to return during last migration. Mexican migration is both an economic and social process. Once migrants are joined by their spouse and children they gradually develop social and economic ties in destination areas; these ties reduce the likelihood of return.

The savings incentive associated with increased last trip duration is strongest for migrants who can convert current foreign earnings into a source of long term income in their place of origin. On the other hand, savings in the form of remittances during the first trip are intended to cover basic needs of the household. Therefore, employment opportunities in community of origin indicate the degree of likelihood that migration is motivated by the need to cover current household expenses as opposed to the simple desire to accumulate savings.

5.2 Future Research

Results from chapter 2 say nothing about the efficiency of environmental policy in stimulating research. Indeed, these results are consistent with (but do not imply) a regulator who chooses standards that are ex-post efficient – that is, efficient for any given state of technology – but not chosen with ex-ante commitments that account for impacts on research incentives (see Requate, 2005b; Innes and Bial, 2002; Innes, 2006). Hence, there is no evidence per se that regulators set tighter standards – vis-à-vis those that are ex-post efficient – in order to spur more innovation, as one might interpret Michael Porter's (1990) famous conjecture to imply.

This observation, as well as the aggregations we make in this study, suggests natural avenues for further inquiry. For example, how do different forms of regulation – tighter standards vs. voluntary pollution reduction programs vs. updated technological regulations – affect innovative effort? And how do different types of innovative effort (more exploratory vs. more derivative) influence and get influenced by environmental standards and regulation? Finally, is there any sense in which regulatory strategy is optimal in inducing and responding to environmental innovation? All of these issues, we believe, merit further study.

Future research from chapter 4 will address the issue of potential effect of unobserved heterogeneity on the hazard rate. Failure to account for heterogeneity may bias the empirical estimates. Also, whether immigrants stay longer when they have access to social assistance. Borjas (1999) finds welfare-receiving immigrants tend to be clustered in certain locations. Moreover, we can tell whether the magnetic effects of welfare (if they affect duration) differ across migrations under different admission categories.

REFERENCES

Angelucci, Manuela (2005) " U.S. Border Enforcement and the Net Inflow of Mexican Illegal Migration" University College London

Anton, Wilma Rose Q., George Deltas, and Madhu Khanna (2004) “Incentives for environmental self-regulation and implications for environmental performance.” Journal of Environmental Economics and Management 48(2004) 632-654

Arellano M., and S. Bond (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations.” Review of Economic Studies 58 (1991): 277-297.

Arellano, M. and O. Bover (1995) “Another look at the instrumental variable estimation of error-components models.” Journal of Econometrics 68 (1995): 29-51.

Arora, S. and T. Cason (1996) “Why do firms volunteer to exceed environmental regulations? Understanding Participation in EPA’s 33/50 program.” Land Economics 72 (1996): 413-32.

Arora, S. and S. Gangopadhyay (1995) “Towards a Theoretical Model of Voluntary Overcompliance.” Journal of Economic Behavior and Organization 28(1995): 289-309

Baltagi, B. H. (1995) Econometric analysis of panel data. John Wiley and Sons, Chichester, 1995.

Bartel, A.P. (1989) "Where do the New Unites States Immigrants Live?" Journal of Labor Economics 7(1989): 371-391.

Bauer, Thomas, Gil S. Epstein and Ira N. Gang. (2003) “Herd Effects or Migration Networks? The Location Choice of Mexican Immigrants in the US.” September 19, 2003

Bean, F.D., E.E. Telles, and B.L. Lowell (1987) "Undocumented Migration to the United States: Perceptions and Evidence." Population and Development Review 13(1987):671-690

Blundell, R. and S. Bond (1998) "Initial conditions and moment restrictions in dynamic panel data models." Journal of Econometrics 87(1998): 115-143.

Borjas, George. J. (1995) "The Economic Benefits from Immigration." Journal of Economic Perspectives 9(1995):3-22.

Borjas, George J. (1987) "Self-Selection and Earnings of Immigrants." American Economic Review 77(1987):531-553.

Borjas, George J. (1999) "Immigration and Welfare Magnets." Journal of Labor Economics 17(4):607-637.

Borjas, George J. and Marta Tienda (1993) "The Employment and Wages of Legalized Immigrants" International Migration Review 27(1993):712-747.

Brunnermeier, S. and M. Cohen (2003) "Determinants of environmental innovation in US manufacturing industries." Journal of Environmental Economics and Management 45(2003) 278-293.

Cameron, C., P. Trivedi, F. Milne, and J. Piggott (1988) "A microeconometric model of the demand for health care and health insurance in Australia." Review of Economic Studies 55(1988): 85-106.

Cameron, C. and Trivedi, P. (1998) Regression analysis of count data. Cambridge University Press, Cambridge, 1998.

Carrión-Flores, Carmen and R. Innes (2006) "Environmental Innovation and Environmental Policy: An Empirical Test of Bi-Directional Effects." Working Paper University of Arizona.

Chiquiar, Daniel and Gordon H. Hanson (2006) "International Migration, Self-Selection, and the distribution of wages: Evidence from Mexico and the United States." NBER working papers series. WP # 9242. October, 2002.

Chiswick, Barry R and Paul W. Miller (2004) "Where Immigrants Settle in the United States" Journal of Comparative Policy Analysis 6(2004): 185-197.

Cuecuecha, Alfredo. (2003) "The Educational Characteristics of the Immigrants from Mexico to the US" Working Paper. Boston University.

Decker, C. and C. Pope (2006) “Adherence to environmental law: the strategic complementarities of compliance decisions.” Quarterly Review of Economics and Finance 45 (2006): 641-61.

Deily, M. and W. Gray (2007) “Agency structure and firm culture: OSHA, EPA, and the steel industry.” Journal of Law, Economics, & Organization (2007), in press.

Donato, Katharine M., Jorge Durand and Douglas S. Massey (1992) "Stemming The Tide? Assessing the Deterrent Effects of the Immigration Reform and Control Act" Demography 29(1992):139-157.

Durand, J. and D.S. Massey (1992) "Mexican Migration to the United States: A Critical Review." Latin American Research Review 27(1992):3-42.

Dustmann, Christian (2003) "Return Migration, wage differentials, and the optimal migration duration" European Economic Review 47(2003): 353-369.

Friedberg, Rachel M. and Jennifer Hunt (1995) "The Impact on Host Country Wages, Employment and Growth" The Journal of Economic Perspectives. 9(1995):23-44.

Fischer, Carolyn, Ian Parry and William Pizer (2003) ‘’Instrument choice for environmental protection when technological innovation is endogenous.’’ Journal of Environmental Economics and Management 45(2003) 523-545

Garcia, C. (2005) "Buscando Trabajo: Social networking among immigrants from Mexico to the United States" Hispanic Journal of Behavioral Sciences. 27(1):3-22

Gourieroux, C., A. Monfort, and A. Trognon (1984) "Pseudo maximum likelihood methods: Applications to Poisson models." Econometrica 52 (1984): 701-720.

Gray, W. and M. Deily. "Compliance and enforcement: Air pollution regulation in the US steel industry." Journal of Environmental Economics and Management 31 (1996): 96-111.

Greene, William H. (2002) Econometric Analysis. 5 th Edition. 790-801.

Grogger, J. (1990) "A simple test for exogeneity in probit, logit, and poisson regression models." Economics Letters 33 (1990): 329-332.

Hansen, L. (1982) "Large sample properties of generalized method of moments estimators," Econometrica 50 (1982): 1029-1054.

Hansen, L., and K. Singleton (1982) "Generalized instrumental variables estimation of nonlinear rational expectations models." Econometrica 50 (1982): 1269-1286.

Hanson, Gordon H. and Antonio Spilimbergo. (1999) "Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the US.-Mexico Border." American Economic Review. 89(1999):1337-1357

Innes, R., and J. Bial. (2002) "Inducing innovation in the environmental technology of oligopolistic firms." Journal of Industrial Economics 50(2002): 265-287.

Innes, R. (2006) "Environmental Policy, R&D and the Porter Hypothesis in a Model of Stochastic Invention and Differentiated Product Competition by Domestic and Foreign Firms." Working Paper, University of Arizona, 2006.

Jaeger, David A. (2004) "Local Labor Markets, Admission Categories, and Immigrant Location Choice." Working Paper IZA.

Jaffe, A., and K. Palmer. (1997) “Environmental regulation and innovation: A panel data study.” The Review of Economics and Statistics. 4 (1997): 610-619.

Jaffe, A. and R. Stavins. (2002) “Environmental policy and technological change” Environmental and Resource Economics 22 (2002): 41-69

Kalbfleish, J.D. and R.L. Prentice (1980) The Statistical Analysis of Failure Time Data. New York: John Wiley & Sons. 1980

Khanna, Madhu and Lisa A. Damon. (1999) “EPA’s Voluntary 33/50 Program: Impact on Toxic Releases and Economic Performance of Firms.” Journal of Environmental Economics and Management 37 (1999) 1-25

Kiefer, Nicholas M.(1988) “Economic Duration Data and Hazard Functions.” Journal of Economic Literature. 26(1988):646-679.

Lancaster, T. (1979), “Econometric Methods for the Duration of Unemployment” Econometrica 47(1979): 939-956.

Lanjouw, J. and A. Mody. (1996) “Innovation and the international diffusion of environmentally responsive technology.” Research Policy. 25 (1996): 549-571.

Laplante, B and P. Rilstone (1996) ‘’Environmental inspections and emissions of the pulp and paper industry in Quebec ‘’ Journal of Environmental Economics and Management 31(1996) 19-36

Lindstrom, David. (1996) "Economic Opportunity in Mexico and Return Migration from the United States" Demography. 33(1996):357-374

Levin, A., C. Lin, and C. Chu. (2002) “Unit root tests in panel data: Asymptotic and finite sample properties.” Journal of Econometrics 108 (2002):1-24.

Magat, W., and W. Viscusi. (1990) “Effectiveness of the EPAs regulatory enforcement: The case of industrial effluent standards.” Journal of Law and Economics 33 (1990): 331-360.

Managi, S., J. Opaluch, D. Jin, and T. Grigalunas. (2005) “Environmental regulations and technological change in the offshore oil and gas industry.” Land Economics 81 (2005): 303-319.

Massey, D.S., J. Durand, and N. Malone. (2002). Beyond Smoke and Mirrors. New York, Russell Sage Foundation.

Massey, D.S., R. Alarcon, J. Durand, and H. Gonzalez. (1987). Return to Aztlan: The Social Process of International Migration from Western Mexico. Berkeley: University of California Press.

Meyer, Bruce. (1990) "Unemployment Insurance and Unemployment spells." Econometrica. 58(1990):757-783.

Milliman, S.R., and R. Prince (1989) ‘’Firm Incentives to Promote Technological change in pollution control.’’ Journal of Environmental Economics and Management. 17(1989) 247-265

Mullahy, J. (1997) “Instrumental-variable estimation of count data models: Applications to models of cigarette smoking behavior.” Review of Economic and Statistics 79 (1997): 586-593.

Munshi, Kaivan. (2003) "Networks in the Modern Economy: Mexican Migrants in the US. Labor Market." The Quarterly Journal of Economics. May(2003):549-599.

Nadeau L (1997) ‘’ EPA effectiveness at reducing the duration of plant-level noncompliance.’’ Journal of Environmental Economics and Management 34(1997) 54-78

Newey, W., and K. West. (1987) “Hypothesis testing with efficient method of moments estimation.” International Economic Review 28 (1987): 777-787.

Popp, D. (2002). “Induced Innovation and energy prices” American Economic Review 92 (2002):160-180.

Porter, M. (1990) The competitive advantage of nations. MacMillan, London, 1990.

Requate, T. (2005) “Dynamic incentives by environmental policy instruments – A survey.” Ecological Economics 54 (2005a): 175-95.

Requate, T. (2005) “Timing and commitment of environmental policy, adoption of new technology, and repercussions on R&D.” Environmental and Resource Economics 31 (2005b): 175-99.

Sam, A. and R. Innes (2006) “Voluntary Pollution Reductions and the Environmental Law: An Empirical Study of the 33/50 Program” forthcoming Journal of Law and Economics

Sjaastad, Larry A. (1962) "The Costs and Returns of Human Migration." Journal of Political Economy 70(1962):80-93

Vernez, G. and D. Ronfelt. (1991) “The Current Situation in Mexican Immigration.” Science 251(1991):1189-1193

Videras, J. and A. Albernini. (2000) “The Appeal of Voluntary Environmental Programs: Which Firms Participate and Why.” Contemporary Economic Policy 18(2000): 449-461

Waldorf, Brigitte. (1995) "Determinants of International Return Migration Intentions." Professional Geography 47(1995):125-136

Windmeijer, F., and J. Santos Silva (1997)"Endogeneity in count data models: An application to demand for health care." Journal of Applied Econometrics 12 (1997): 281-294.

Windmeijer, F. (2002) ExpEnd, A Gauss Programme for Non-Linear GMM Estimation of Exponential Models with Endogenous Regressors for Cross Section and Panel Data. Institute for Fiscal Studies (2002).

Windmeijer. F. (2000) "A finite sample correction for the variance of linear two-step GMM estimators." IFS working paper W00/19, Institute for Fiscal Studies (2000).