THREE ESSAYS IN INTERNATIONAL TRADE: ANALYSIS OF EXPORT DECISIONS AND OFFSHORING

By

Xuan Wei

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food and Resource Economics – Doctor of Philosophy
Economics – Doctor of Philosophy

2013

ABSTRACT

THREE ESSAYS IN INTERNATIONAL TRADE:
ANALYSIS OF EXPORT DECISIONS AND OFFSHORING

By

Xuan Wei

Trade liberalization through reductions in trade barriers by bilateral and multilateral agreements boosts individual firm engagement in international trade. The importance of firm decisions in exploring a new foreign market has been increasingly recognized in terms of understanding the causes and consequences of aggregate trade flows. Meanwhile, strategic participation in international trade in turn stimulates the process of global integration, and impact the diversification of international trade from simple trading goods to more complicated trade in service through channels such as offshoring. These new features of international trade cannot be fully explained by pre-existing trade theories.

This dissertation, consisting of three essays focused on the reciprocal relationship between international trade and firm behaviors, provides additional empirical evidence on the interactions between these two and contributes to the development of theoretical research along these lines.

In the first essay, a theoretical model is developed to analyze how an individual firm may reduce or eliminate the uncertainty of trade compliance costs associated of entering a foreign market by paying for information to making export decisions. I extend the heterogeneous firm model of Melitz (2003) to show that in the presence of uncertain compliance costs and non-zero information costs, average profits and productivity differences between exporting and non-exporting firms are reduced.

The second essay investigates the effect of information costs and compliance costs on firm decisions to export as well as how much to export through a hurdle model approach. A bilateral trade flow data at SITC4-digit industry level from 1991-2000 are used to

approximate the export and value decisions of heterogeneous firms. Results show the effect of fixed export costs is twofold: Information costs decreases the probability of export by about five to six percentage points in the first stage. Once the export decision is made, firms that paid information costs in the first stage tend to export more in the second stage to compensate such costs. On the other hand, paying compliance costs decreases the probability of export by about 36 percentage points. Compliance costs are more prohibitive in the subsequent export value decision.

The third essay generalizes the Grossman and Rossi-Hansberg (2008) offshoring model to include numerous tasks/skill levels. This generalization allows a possible and direct linkage between the theoretical task offshoring model and occupational data that can be aggregated from the CPSMORG (Current Population Survey Merged Outgoing Rotation Groups) data from year 1983 to 2011. Empirical investigation of the effect of offshoirng on occupational employment for the ten major occupational groups (at 2-digit SOC level) in the U.S. labor market is conducted by estimating their offshoring cost functions using a non-parametric monotonic cubic spline interpolation method. Five relatively offshorable occupational groups are identified from the estimated offshoring cost functions.

The number of jobs offshored and the offshoring percentage for the five relatively offshorable occupational groups under three scenarios are calculated under NLS (non-linear least squares) method by attaching a cubic offshoring cost functional form to all five groups. Results show production occupations are the most offshorable while sales and related occupations are the least offshorable among all five groups under all three scenarios. Offshoring percentage for production occupations has been increasing in both pre- and post-2000 periods while the offshoring percentages for professional and related occupations, and management, business, and financial operations occupations have been decreasing over time.

For
my mother Yuefen Wei and my father Jinlin Wu

ACKNOWLEDGEMENTS

This dissertation is an outcome of training myself to be an independent scholar with abilities of identifying a question, proposing a method and resolving a problem. During this process, I have been endowed with tremendous support and assistance from many people. Without them this dissertation would have not been possible.

First of all, I would like to express the appreciation to my Major Professor and Dissertation Committee Chair Dr. Suzanne Thornsbury. I especially appreciate her willingness of continuously serving as my Dissertation Committee Chair even after she left MSU in 2011. To me, she is not only an academic mentor throughout my doctoral studies, but also a great friend to consult with concerns and problems as well as to share moments in my life. She is particularly good at directing students to think through the fundamental issues of a research. Under her guidance, I learned how to refine research questions by sorting out different phenomena, and then turn a complicated issue into a tangible problem.

I am grateful to Dr. David Schweikhardt (Co-chair), who is willing to serve as my Dissertation Committee Co-chair from a middle point, and other two committee members, Dr. Susan Chun Zhu and Dr. Jeffery Wooldridge for their deep engagement and valuable comments to improve my dissertation. I am so fortunate to have such a balanced committee, each with strengths in different aspects. In particular, I would like to extend my gratitude to Dr. Zhu for her supervision of shaping my 3rd essay idea and developing it to a dissertation chapter. I thank Dr. Wooldridge for his tremendous contributions on choice of econometric methods and models, comparisons between model specifications and interpretation of empirical results for both my 2nd and 3rd essays. I would like to thank all my committee members for their tolerance of allowing me to work from distance and maintain contact with them by emails and phone calls so that I can stay with my family in Madison, WI.

I would like to thank the Elton R. Smith Endowment funding and the College of Agriculture and Natural Recourses Dissertation Competition Fellowship for additional financial support of my dissertation research.

A special thank goes to Dr. Scott Swinton for his strong support in many aspects including funding arrangement and academic advice. I also thank Debbie Conway, the Department Secretary for her always-ready-to-help attitude and effective assistance in dealing with administrative process when I was off campus.

My acknowledgements would be incomplete if I don't mention help and support from my fellow colleagues, faculty and staff members at MSU as well as friends at UW-Madison. I am especially grateful to, Dr. Sandra Batie, Mollie Woods, Lori Jean, Ge Bai, Shan Ma, Fang Xia, Min Chen, Hui Wang, Li Cheng, Chenguang Wang, Chaoran Hu, Tim Komarek, Christina Plerhoples, Yuliang Wang and Jinny Chen.

In addition, a thank you to Dr. Tadashi Yamada, Dr. Nobuyuki Hanaki and Dr. Mamoru Kaneko at Univeristy of Tsukuba (Tsukuba Daigaku, Japan) who introduced me to the world of economics and encouraged me to pursue advanced studies in economics.

My mother Yuefen Wei and father Jinlin Wu receive my deepest gratitude and love for their unconditional support and respect of every choice I make for myself. Without their constant encouragement, I wouldn't be where I am today. As I am struggling through the revision of my dissertation, my mother is fighting against her sickness. Her positive attitude inspires me and I believe that we together will win the battle.

I also thank my sister Wei Wu and brother-in-law Weizhong Zhu for their financial support, for helping me fulfill the responsibility of a dutiful daughter and grand-daughter as I am not able to. Sadly, my grandma was not able to wait till the day coming to share the joy with me, she passed away on February 10, 2013.

Finally, thanks to my husband Xianwei Meng, my daughter Elle Q. W. Meng and my son William Q. W. Meng. I am so lucky to meet Xianwei, my soul mate at MSU. He is also such a great partner to discuss my initial research ideas with, to debate about different theories, and to provide a proofread of my mathematical calculation and derivation. During the most stressful days of dissertation writing, he always has the magic power to calm me down and rekindle my hope. Thank you for your faith in me and always being there for me. Without your company and support, I truly cannot do this.

The first fruit of our marriage, Elle, my sweet daughter, was born on June 27, 2011. As a mom, I beg Elle's forgiveness for not being able to stay with her and witness every step of her growth. It was a difficult decision to send Elle back to China so that I can focus on my dissertation and complete my degree. In this sense, I am deeply indebted to my parents, especially my mother for taking care of Elle with love. For those days that I have missed with Elle, I wish I could make up for it with the rest of my life.

My second little one William was born on April 9, 2013. Taking care of a newborn and revising the dissertation at the same time are frustrating sometimes, but the joy is overwhelming. I am glad that Xianwei shared the responsibility to looking after

William. I would never forget the scenes that William sleeps comfortably on his daddy's chest in the first month.

My dear mom and my little sweetie pies, because of you, every day is full of meaning to me. Because of you, I am stronger and will be even stronger.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Despite significant reductions in tariff barriers and revolutionary advances in transportation and communication technologies, individual firms still face various trade costs in order to enter a foreign market. In 2004, ad-valorem tax equivalent trade costs for industrialized countries were estimated to be 170 percent of producer price (Anderson and van Wincoop, 2004). Included are impacts from tariffs (less than 5 percent), nontariff barriers (8 percent), and information costs (6 percent), which implies these three sources were almost equal impediments to trade on an aggregated basis. As tariffs are reduced through bilateral and multilateral agreements, concern over the substitution of non-tariff trade barriers has increased. Among them, increased compliance costs (costs which are necessary to conform with the regulations governing market access) are an obvious consequence from the proliferation of non-tariff barriers, which are uncertain prior to export or collection of information which is costly for an individual firm.

In the first essay, the heterogeneous firm model of Melitz (2003) is extended to analyze how an individual firm may reduce or eliminate the uncertainty of compliance costs by paying for the information cost prior to making decisions to export. Firms' ability to eliminate uncertainty over compliance costs depends on their undertaking the information cost necessary to learn about regulations. Results suggest that in the presence of uncertain compliance costs and non-zero information cost, average profits and productivity differences between exporting and non-exporting firms are reduced.

Two separate scenarios are discussed where the presence of positive information cost will suppress the likelihood of exporting. In the first scenario, firms capable of exporting in the absence of information cost will be limited to the domestic market due to prior overestimation of compliance costs and lack of incentive to pay the information cost. As a consequence, firms with high productivity trapped in the domestic market will increase the overall productivity and profits of domestic firms, (i.e., non-exporting firms). In the second scenario, firms with low productivity are overoptimistic to pay the information cost but not able to export due to prior underestimation of compliance costs. Exit of some low-productivity firms from the domestic market due to the loss of information cost will also in turn increase the overall productivity and profits of domestic firms, (i.e., non-exporting firms).

Partially inspired by the theoretical prediction in the first essay, the second essay attempts to seek empirical evidence of the existence of information costs and compliance costs using real data. The second essay uses a hurdle model approach to separate the effect of information costs and compliance costs and empirically examines the effect of the presence of positive information costs and compliance costs on (a) the likelihood of an individual firm's export decision in the first stage, and (b) the export value in the second stage after the export decision is made.

To capture the two-stage export decision process of an individual firm, a hurdle model approach is employed. In particular, two model specifications, lognormal hurdle model and exponential type II Tobit (ET2T) model (or Heckman method) are adopted and results are compared between these two models. Two indicators are created to infer and separate the existence of information costs and compliance costs based on an individual firm's export status in previous years. Under some theoretical restrictions, a bilateral trade flow data at SITC4-digit industry level from 1991-2000 are used to approximate the export and value decisions of heterogeneous firms.

A statistical χ^2 test is strongly in favor of the exponential type II Tobit (ET2T) model (correlations between the two-stage decisions are assumed) against the lognormal hurdle model (independence between the two-stage decisions are assumed). Results from the exponential type II Tobit (ET2T) model indicate the effect of fixed export costs is two-fold. Information costs are crucial in determining whether or not to export. It decreases the probability of export by about five to six percentage points in the first stage. Once the export decision is made, firms paying information costs tend to export in the second stage to cover such costs. On the other hand, paying compliance costs decreases the probability of export by about 36 percentage points. Compliance costs are more prohibitive in the subsequent export value decision.

If the first two essays directly evaluate the existence of trade barriers on firms' participation in international trade, then the third essay investigates the consequences of firms' participation through offshoring in international trade as offshoring has spread rapidly from the jobs of blue-collar workers in manufacturing sectors to those of white-collar workers in service sectors. Workers in all sectors in most developed countries, including the United States became more concerned about the security of their jobs as the global economy continued to integrate.

The third essay generalizes the Grossman and Rossi-Hansberg (2008) offshoring model to include numerous tasks/skill levels (tasks correspond to specific occupations in the empirical framework) and then empirically investigate the effect of offshoirng on occupational employment for major occupational groups (at 2-digit SOC level) in the U.S. labor market by (a) estimating the offshoring cost functions for the ten major occupational groups and identifying relatively offshorable occupational groups based on estimated offshoring costs; (b) calculating the number of jobs offshored and offshoring percentage for relatively offshorable occupational groups.

This research uses the CPSMORG (Current Population Survey Merged Outgoing Rotation Groups) data from year 1983 to 2011 aggregated at the occupational level. A non-parametric monotonic cubic spline interpolation method is introduced to approximate the offshoring cost functions for the ten major occupational groups because standard parametric methods cannot directly estimate functions. Estimated offshoring costs demonstrate that among the ten occupational groups, groups involved with more impersonal and/or routine-tasks have relatively lower offshoring costs in comparison to groups involved more personal and/or non-routine manual tasks. Included in the former group are production occupations and office and administrative support occupations. Occupations in the latter group include farming, fishing, construction, extraction, repair, and transportation. This finding is consistent with the initial hypotheses and with Blinder and Krueger (2009) who examined offshorability in major occupational groups based on a telephone survey.

Motivated by the practical issue of difficulty in obtaining a time-variant offshoring/offshorability index faced by majority empirical studies interested in identifying the effect of offshoring, this research calculates the number of jobs offshored as well as the offshoring percentage by occupation for the five relatively offshorable occupations over the sample period under three different scenarios. My calculation indicates that production occupations are most offshorable among all five offshorable occupational groups in all three scenarios. In addition offshoring percentage for production occupations has been increasing in both pre- and post-2000 periods while the offshoring percentage for professional and related occupations, and management, business, and financial operations occupations has been decreasing over time. The model's calculation also provide time-variant offshoring indices for more than 300 major U.S. detailed occupations in these five relatively offshorable groups that can be employed in other empirical studies.

REFERENCES

REFERENCES

Anderson, J. E. and E. van Wincoop (2004). Trade costs. *Journal of Economic Literature* 42 (3), 691–751.

Blinder, A. S. and A. B. Krueger (2013). Alternative measures of offshorability: A survey approach. *Journal of Labor Economics 31* (2), S97–S128.

Grossman, G. M. and E. Rossi-Hansberg (2008). Trading tasks: A simple theory of offshoring. *American Economic Review 98* (5), 1978–97.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71* (6), 1695–725.

CHAPTER 1: FIRM LEVEL EXPORT DECISIONS: THE ROLE OF INFORMATION COST[1]

1.1 Introduction

This paper investigates how individual firms eliminate or reduce uncertainty over compliance costs by paying for information prior to making the decision to export. I explicitly incorporate such costs into the heterogeneous firm model of Melitz (2003). Difficulties in accessing information about regulations or exporting procedures could lead to over- or under-estimation of compliance costs and thus limit export competitiveness well beyond the content of regulations themselves. Even though regulations and standards are sometimes publicly accessible through official web sites, it is not an easy or costless task for firms to go through the tedious, and often obscure, documentation to extract the specific information they need. Anderson and van Wincoop (2004) estimated the ad-valorem tax equivalent trade costs for industrialized countries and showed that tariffs, nontariff barriers, and information costs were almost equal impediments to trade on an aggregated basis.

Increased compliance costs are an obvious consequence from proliferation of non-tariff barriers (Thornsbury et al., 2004). A survey across exporting firms in three diverse industries (telecommunications equipment, dairy products, automotive components) found that many firms had difficulty determining compliance costs *ex ante* with part of the uncertainty related to conformity assessment (for example, inconsistent product evaluation by assessment bodies) (OECD, 2000). Firms had to make decisions about modifications to meet anticipated foreign requirements without full information. However, compliance costs are

[1] For citation purposes please refer to the article published earlier: Wei, X. and S. Thornsbury (2012). Firm Level Export Decisions: the Role of Information Cost. *Economics Letters* 116, 487-90.

usually analyzed as if they were certain prior to an export decision (Markus et al., 2001). Information costs as well as the uncertainty of compliance costs due to lack of information are not considered and excluded from this line of research.

Melitz (2003) developed a heterogeneous firm model to examine the effect of compliance costs to export markets on different types of firms, but information cost and uncertainty over compliance costs were assumed to be nonexistent. This work extends the Melitz model by explicitly considering these two elements: 1) collecting information is not free and will be sunk prior to the export decision; 2) compliance cost for each individual firm is uncertain prior to information collection. Analogous to one-time compliance costs, information cost is a fixed expense necessary for exporting.

1.2. A Conceptual Model

In this section, the model examines the effect of non-zero information cost on the likelihood that a firm will export. This model partitions the concept of export cost $\left(f_{ex}\right)$ in the heterogeneous firm model of Melitz (2003) into fixed information cost $\left(f_{ic}\right)$ and compliance costs $\left(f_{cc}\right)$ in the open economy.[2]

Information cost $\left(f_{ic}\right)$ is incurred by a firm in order to participate in foreign markets. Investment in information collection must be undertaken in a period before the real fixed compliance costs $\left(f_{cc}\right)$ is revealed; thereafter information costs are sunk (Figure 1).[3]

[2] I retain the Melitz assumptions, derivations and theoretical framework when the economy is closed.

[3] The timing is not altered whether a firm collects all necessary information in-house or hires

Prior to paying the information cost $\left(f_{ic} \right)$, firms only have some initial belief that compliance costs f_{cc} falls within a certain range. This *ex-ante* belief about compliance costs is denoted as a random variable B which follows a distribution $h(B)$ on $[0,\infty)$ with *cdf* $H(B)$.[4]

For the purpose of comparison, the Melitz modeling strategy of per-period representation is adopted for the remaining discussion. If firms are assumed to not discount the future (i.e., discount factor $\rho = 1$), then firms are indifferent between paying the one time fixed export cost f_{ex} or paying a proportion of this cost in every period. Similarly, firms are indifferent between paying f_{ic} or paying a per-period information cost $f_i \left(= \delta f_{ic}\right)$ and indifferent between paying B or paying a per-period anticipated compliance costs $\beta\left(= \delta B\right)$.[5] Provided that the real compliance costs $\left(f_{cc} \right)$ is revealed after paying the information cost but prior to export decision, the actual per-period compliance

---

a broker to provide such service. See footnote 7 for detailed explanation.

[4] To be consistent with the Melitz model, firm heterogeneity comes solely from productivity differences and all firms share the same distribution of beliefs about compliance costs. The model can be extended to account for heterogeneous belief distributions based on different firm types by duplicating the current analysis for each type of firm. For example, to obtain unique productivity cutoffs, it is reasonable to assume beliefs of high productivity firms are first-order stochastically dominated by those of low productivity firms, i.e., the higher the productivity, the lower the firm's belief about compliance costs. In this case all analytical results remain valid.

[5] The anticipated per-period compliance costs $\beta = \delta \mathrm{B}$ has the same distribution as B.

costs paid by a firm is equal to $f_c \left(= \delta f_{cc}\right)$. There is no randomness in realization of the actual per-period compliance costs f_c.

Assuming a continuum of firms with different productivity levels indexed by φ, where φ is distributed as $g(\varphi)$ with *cdf* $G(\varphi)$,[6] only firms with productivity $\varphi \geq \varphi^*$ remain in the industry. Hence by construction, a firm seeking to export must have a productivity level no less than φ^*.

With the existence of information cost, the profit from domestic sales in every period remains $\pi_d(\varphi) = \dfrac{r_d(\varphi)}{\sigma} - f$ but profit from export is modified to $\pi_x(\varphi) = \dfrac{r_x(\varphi)}{\sigma} - f_i - f_c$, where $r_d(\varphi)$, $r_x(\varphi)$ are domestic and export revenues, σ is the constant elasticity of substitution, f is the fixed cost for production,[7] and f_i and f_c are per-period information cost and per-period compliance costs respectively.

---

[6] Following Melitz (2003), all firms share the same fixed cost $f > 0$. Labor *l,* the only factor required for production, is a linear function of output $l = f + q/\varphi$. This implies marginal cost is $1/\varphi$ and higher φ indicates a higher productivity level.

[7] See footnote 5 for specification of production function.

Among firms staying in the industry, the marginal firm which pays information cost f_i and is actually exporting must have productivity level φ_x^* satisfying

$$\frac{r_x(\varphi_x^*)}{\sigma} - f_i - f_c = 0 \qquad \text{(zero profit condition).} \qquad (1.1)$$

By definition, per-period export cost should equal the information cost plus the compliance costs, i.e., $f_x = f_i + f_c$.[8]

Since information about compliance is not costless, the model needs to first consider which firm will undertake these actions to export. A firm pays information cost and earns the expected profit denoted as $I(\varphi)$, which is given by:

[8] After a broker collected export and compliance information for the first firm, marginal cost of providing the information to an additional firm approaches zero. Under a competitive brokerage service assumption, the marginal cost of information collection would then equal the price of information $f_i = 0$. When information cost $f_i = 0$, compliance cost f_c becomes certain, $f_x = f_c$, and results revert to the original Melitz solution. In reality, both brokers and exporting firms consider information proprietary and a potential source of competitive advantage. Small firms often use brokerage services and pay information cost between $f_i = 0$ and $f_i (= \delta f_{ic})$.

$$I(\varphi)=\int_0^{\frac{r_x(\varphi)}{\sigma}-f_i}\left(\frac{r_d(\varphi)}{\sigma}-f+\frac{r_x(\varphi)}{\sigma}-f_i-\beta\right)h(\beta)d\beta$$
$$+\int_{\frac{r_x(\varphi)}{\sigma}-f_i}^{\infty}\left(\frac{r_d(\varphi)}{\sigma}-f-f_i\right)h(\beta)d\beta$$

(1.2)

The first integral is the expected profit from both domestic market $\left(\frac{r_d(\varphi)}{\sigma}-f\right)$ and foreign market $\left(\frac{r_x(\varphi)}{\sigma}-f_i-\beta\right)$ when the anticipated compliance costs β is less than $\frac{r_x(\varphi)}{\sigma}-f_i$ and a firm exports. The second integral is the expected profit from domestic market net information cost when the anticipated compliance costs β is greater than $\frac{r_x(\varphi)}{\sigma}-f_i$ and a firm will not export.

After some simplification,

$$I(\varphi)=\int_0^{\frac{r_x(\varphi)}{\sigma}-f_i}\left(\frac{r_x(\varphi)}{\sigma}-\beta\right)h(\beta)d\beta+\frac{r_d(\varphi)}{\sigma}-f-f_i\,.$$

(1.2)'

The marginal firm, indifferent between paying the information cost to export or staying in the domestic market, is the firm with productivity level $\hat{\varphi}$ satisfying the free entry condition below, such that expected profit from exporting equals domestic profit ,

$$I(\hat{\varphi}) = \pi_d(\hat{\varphi})$$ (free entry condition).

(1.3)

From the above free entry condition, $J(\varphi)$ is denoted as the difference between expected profit from exporting and domestic profit for a firm with productivity level φ,

$$J(\varphi) = I(\varphi) - \pi_d(\varphi) = \int_0^{\frac{r_x(\varphi)}{\sigma} - f_i} \left(\frac{r_x(\varphi)}{\sigma} - \beta \right) h(\beta) d\beta - f_i.$$

(1.4)

$J(\varphi)$ is increasing in φ because

$$J'(\varphi) = \frac{r_x'(\varphi)}{\sigma}\left[H\left(\frac{r_x(\varphi)}{\sigma} - f_i \right) + f_i h\left(\frac{r_x(\varphi)}{\sigma} - f_i \right) \right] \geq 0$$ by $r_x'(\varphi) \geq 0$.

When $\varphi = \varphi_x^*$, after some simplification

$$J\left(\varphi_x^*\right)=\int_0^{f_c}\left(\frac{r_x\left(\varphi_x^*\right)}{\sigma}-f_i-\beta\right)h(\beta)d\beta-\int_{f_c}^{\infty}f_i h(\beta)d\beta \quad . \tag{1.5}$$

By zero profit condition (equation (1.1)), Equation (1.5) collapses to

$$J\left(\varphi_x^*\right)=\int_0^{f_c}\left(f_c-\beta\right)h(\beta)d\beta-f_i\left(1-H\left(f_c\right)\right) \quad . \tag{1.6}$$

Hence, the sign of $J\left(\varphi_x^*\right)$ depends on the difference of the two terms in equation (1.6). The first term is the firm's expected profits from exporting, while the second term is the firm's expected loss from information cost when compliance costs are prohibitive.

1.3. Analytical Results

I distinguish two separate scenarios where the presence of positive information cost will suppress the likelihood of exporting.

Proposition 1. *If firms overestimate the real compliance costs a priori* (*i.e., a high probability that prior belief about the compliance costs is large), then* $\varphi_x^* < \hat{\varphi}$ *.*

Proof:

$$J\left(\varphi_x^*\right) < 0 \qquad \Leftrightarrow$$

$$\frac{1}{1-H\left(f_c\right)}\int_0^{f_c}\left(f_c-\beta\right)h\left(\beta\right)d\beta < f_i$$

$$\Leftrightarrow$$

$$\frac{H\left(f_c\right)}{1-H\left(f_c\right)}\left[f_c-\int_0^{f_c}\beta\frac{h\left(\beta\right)}{H\left(f_c\right)}d\beta\right] < f_i$$

$$\Rightarrow$$

$$\frac{H\left(f_c\right)}{1-H\left(f_c\right)}\left[f_c-\int_0^{f_c}\beta\frac{h\left(\beta\right)}{H\left(f_c\right)}d\beta\right] < \frac{H\left(f_c\right)}{1-H\left(f_c\right)}f_c < f_i$$

In the last line, the first inequality holds because the integral in the bracket is always positive. Hence, to make the second inequality hold requires $H\left(f_c\right)$ *being small* *enough (i.e., there is a very small probability that a firm's prior belief about the compliance costs* (β) *is less than the real compliance costs* (f_c)*, or a firm overestimates the compliance costs). Given* $J(\varphi)$ *is increasing in* φ *, I have* $\varphi_x^* < \hat{\varphi}$ *.*

Q.E.D.

Firms may have little knowledge about the export process (large uncertainty over compliance costs *ex-ante*). In this case, information costs will be greater and fewer firms will have incentive to pay the information cost. A certain number of firms capable of exporting in the absence of information cost, therefore, are trapped in the domestic market due to prior overestimation of compliance costs f_c . In comparison to the export cut-off productivity level φ_x^* (Figure 1.2a) as originally depicted in Melitz (2003), Figure 1.2b shows that existence of information cost increases the threshold productivity level for exporting from φ_x^* to $\hat{\varphi}$. In the presence of information cost, average profit and productivity levels of non-exporters increase, the number of exporters is reduced, and average profit and productivity differences between exporters and non-exporters are reduced.

Proposition 2. *If firms underestimate the real compliance costs a priori (i.e., a high probability that prior belief about the compliance costs is small), then*

$$J\left(\varphi_x^*\right) \geq 0 \text{ and } \varphi_x^* \geq \hat{\varphi}.$$[9]

In this case, some firms behave too optimistically and thus more firms than necessary are paying the information cost but unable to export once compliance costs are known (see Figure 1.2c). Since firms make their export decision after observing f_c that becomes known through paying the information cost, the number of exporters is not altered in this case. I discuss two extreme cases: either $\widehat{\varphi}$ is close enough to φ_x^*, or $\widehat{\varphi}$ is close enough to φ^*. In the first case, firms still export because they are better off by exporting if $\frac{r_x(\varphi)}{\sigma} \geq f_c$.[10] The Average profit and productivity level of exporters are thus decreased because firms with $\widehat{\varphi}$ close enough to φ_x^* earn negative profits by exporting. In the latter case, marginal firms with $\widehat{\varphi}$ close enough to φ^* might have to exit the

[9] Proof of Proposition 2 is parallel to the previous proof with a reverse of the inequality $J\left(\varphi_x^*\right) \geq 0$.

[10] Firms make less negative profits by exporting than remaining in the domestic market.

industry due to the information cost resulting in $\pi_d(\varphi) = \frac{r_d(\varphi)}{\sigma} - f - f_i < 0$.

Exit of low productivity firms increases average profits as well as the productivity level of non-exporters.

I therefore have the following

Proposition 3. *Uncertainty over compliance costs which can be mitigated through positive information cost decreases average profit and productivity gaps between exporters and non-exporters although the number of exporters may or may not change*.

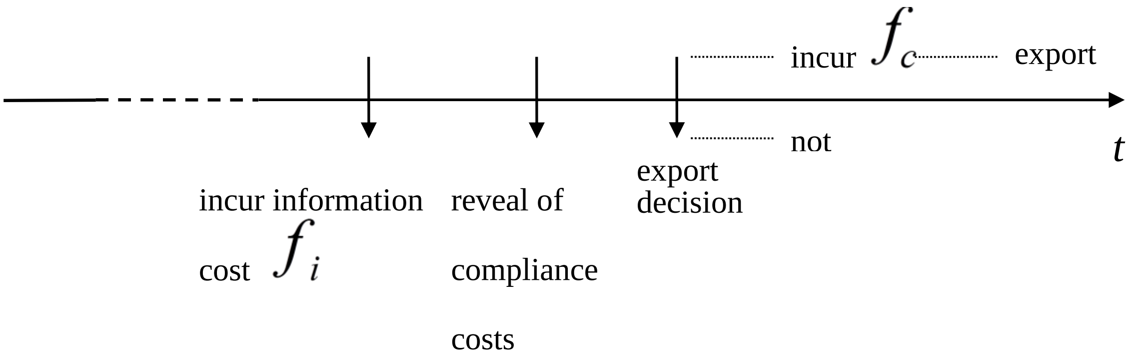
With positive information cost and uncertainty over compliance costs, the difference in productivity and average profit between exporting and non-exporting firms is reduced. This is a new feature not captured in the Melitz standard heterogeneous firm model.

1.4. Conclusions

In the presence of positive information cost, *ex-ante* compliance costs are uncertain for individual firms when making the decision whether or not to export. Firms' ability to eliminate uncertainty over compliance costs depends on undertaking the information cost necessary to learn about regulations. Two scenarios derived from my theoretical model predict that profit and productivity differences between exporters and non-exporters are reduced in industries subject to more stringent regulations where *ex-ante* compliance costs tend to be more uncertain for individual firms without prior costly information collection.

APPENDIX

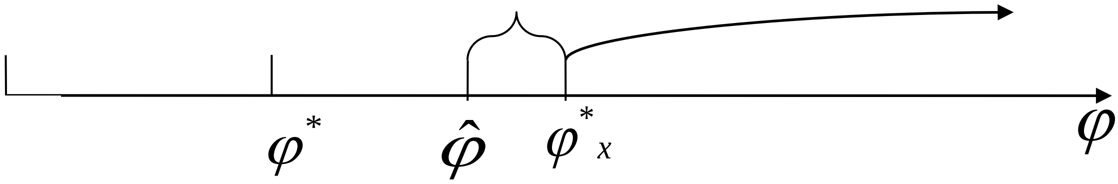
Figure 1.1: Timeline for firm export decisions



Note: All notations in this figure are expressed in per-period representation.

Figure 1.2: Minimum productivity level required for export under alternative depictions of information cost.

1.2a. Firms facing per-period fixed export cost f_x *: Melitz (2003)*



1.2b.Firms facing uncertain compliance costs f_c *prior to paying information cost* f_i

(overestimate)

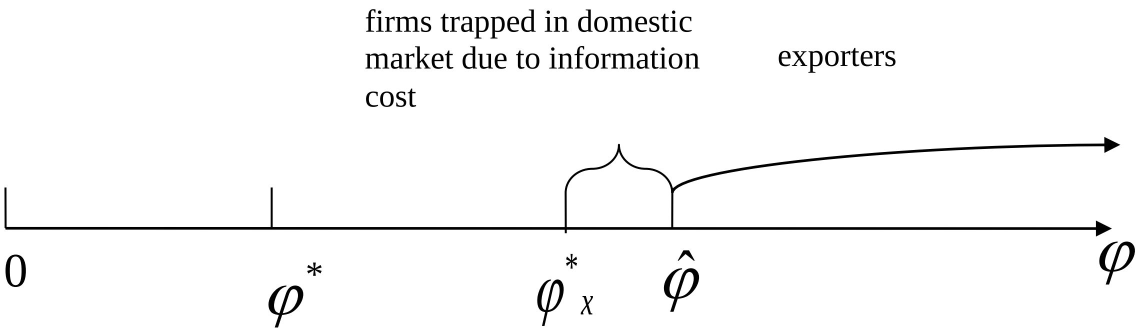
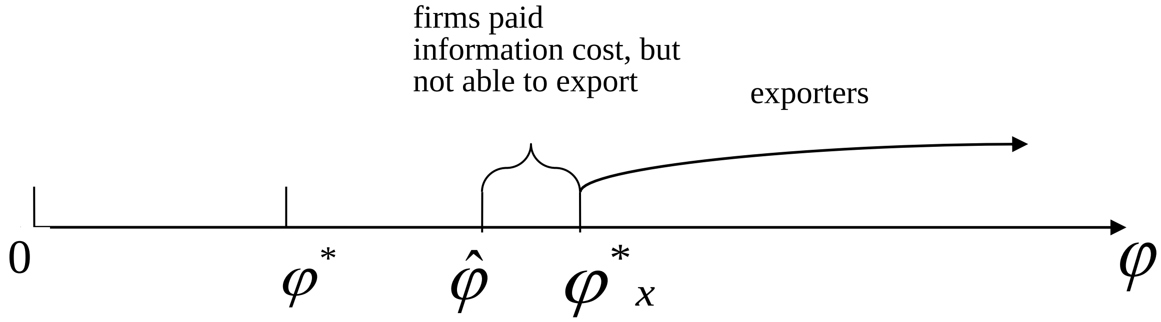


Figure 1.2 (Cont'd)

1.2c.Firms facing uncertain compliance costs f_c *prior to paying information cost* f_i *(underestimate)*



Note: Along the horizontal productivity line, φ is firm productivity. φ^* and φ^*_x define the minimum productivity levels for firms to remain in the industry and to export respectively. $\hat{\varphi}$ defines the productivity level for firms that pay information costs and consider exporting. All notations in this figure are expressed in per-period representation.

REFERENCES

REFERENCES

Anderson, J. E. and E. van Wincoop (2004). Trade costs. *Journal of Economic Literature 42* (3), 691–751.

Maskus, K. E., T. Otsuki and J. S. Wilson. “An empirical framework for analyzing technical regulations and trade,” in *Quantifying the Impact of Technical Barriers to Trade,* ed. Keith E. Maskus and Jone S. Wilson (Ann Arbor : University of Michigan Press, 2001), 29.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71 (6), 1695–725.

Organization for Economic Development and Cooperation. An Assessment of the Costs for International Trade in Meeting Regulatory Requirements. Technical Report. TD/TC/WP(99)8/FINAL, OECD, Paris, 2000.

Thornsbury, S., D. Roberts, and D. Orden (2004). Measurement and political economy of disputed technical regulations. *Journal of agricultural and applied economics* 36 (3), 559-74.

CHAPTER 2: AN EMPIRICAL EXAMINATION OF INFORMATION COSTS AND COMPLIANCE COSTS USING A HURDLE MODEL APPROACH

2. 1. Introduction

Upon export to a new foreign market, firms seeking to export face per-unit export costs such as tariff and transport costs, but also face fixed costs that do not vary with export volume.[1] Two typical fixed export costs faced by a new exporting firm are i) Investment in research to understand the regulations and standards in a potential foreign market (herein called information costs); ii) product redesign for a specific market, establishing new processes or procedures to comply with foreign regulations and standards (herein called compliance costs).[2]

Several studies have focused on using firm level data to validate the existence of fixed export costs and quantify the effect on firm participation in foreign markets in both the extensive margin (i.e., the number of exporting firms) and intensive margin (i.e., volume of each exporting firm). For Columbian manufacturing firms, Roberts and Tybout (1997) found that (a) fixed export costs significantly reduce the probability of a firm exporting and (b) a firm with prior experience is far more likely to export than a firm that has never exported. Bernard and Jenson (2001) showed that entry costs to foreign markets are substantial for U.S. manufacturing firms, and firms are increasingly likely to export in consecutive years. Das et al. (2007) quantified the fixed entry costs for three Colombian manufacturing industries

[1] In some international trade literature, fixed export costs are called fixed entry costs, sunk costs or sunk entry costs.
[2] Some compliance costs can be repeatedly occurring for continuous quality control and testing certification even if a firm keeps exporting to the same market.

(basic chemicals, leather products, and knitted fabrics). Their results indicated that average fixed entry costs to foreign markets were similar across the three manufacturing subsectors between 1981 and 1991, but were lower for large producers (e.g., $402,000 for knitting mills) relative to small producers (e.g., $412,000 for knitting mills). These studies estimated how the existence of fixed export costs affected the extensive margin of trade flow (i.e., how the existence of fixed export costs affect the likelihood of an individual firm to export). Helpman et al. (2008) estimated the effect of fixed export costs on both extensive and intensive margins (i.e., the number of exporting firms and the trade volume per exporting firm respectively) in a theoretical model accounting for both decisions (self-selection of firms into the export markets and firms' export volume) where fixed export costs were assumed to be known.

Two major assumptions in the existing studies that examining the effect of fixed export costs have not been addressed. First, fixed export costs (including the theoretical heterogeneous firm model of Melitz, 2003) are treated as certain when firms make export decisions (Figure 2.1.1). A second issue that most current empirical studies in this field have not addressed is differentiation among export destinations. Compliance costs are often uncertain due to a lack of information prior to the decision of export (Figure 2.1.2). In the presence of positive information costs and uncertainty of compliance costs, the number of exporting firms is smaller than that reported by Melitz (2003) where fixed export costs are assumed to be certain. When compliance costs are uncertain, some firms capable of exporting may overestimate compliance costs without collecting information and thus are trapped within the domestic market.

An individual firm's sequential decision to export can be described by a standard hurdle model often used in empirical settings. In the first stage, an individual firm determines whether or not to export to a foreign market providing total fixed export costs are known. Once the export decision is made, the individual firm determines how much to export in the second stage. However, when there are positive information costs and compliance costs that are not revealed until after an individual firm pays the information costs the first stage export decision is contingent on previous export status, i.e., whether a firm is new to this specific foreign market in period t so that collecting information is required prior to export.

Blanes-Cristóbal (2008) empirically tested and confirmed the differences of fixed export costs across markets, but they only divided export destinations into three general market areas: EU, OECD and ROW. Morales et al. (2011) developed a structural model and estimated the fixed entry cost for chemical and chemical products manufacturing firms using Chilean data. Their results show country-specific entry costs increases significantly as the destination country is farther away and less similar from the exporting country. Simply considering the export decision as a single choice of whether or not to export regardless of destination obscures important market specificity characteristics that define fixed export costs. Although a firm's status as an exporter tends to be persistent, an exporting firm may frequently enter a new foreign market or exit a current export market. Selection of destination countries depends not only on similarities between a new export market and the firm's home country, but also on similarities between a new market and previous destinations (Morales et al., 2011). Thus, a firm's export decision should consider the question of whether or not to export to a specific foreign market because information costs and compliance costs are market specific and can vary substantially depending on (a) which foreign market is selected for expansion and (b) which export market is already served by the firm. It is thus necessary to differentiate export destinations to clarify fixed export costs as a part of quantifying these costs on firm export decisions.

In this paper, the effect of information costs and compliance costs are decomposed to empirically examine how their presence may affect (a) the likelihood of a positive individual firm export decision in the first stage, and (b) the export value in the second stage after the export decision is made in a hurdle model. Panel bilateral trade flow data at SITC4-digit industry level from 1991-2000 are used to approximate the export decision and export value decision of heterogeneous firms. Two indicators are created to infer the existence of information costs and compliance costs based on export status in previous years, as neither information costs nor compliance costs are directly observed in the data. In addition, export destinations are distinguished when examining the effect of information costs and compliance costs. Market heterogeneity provides the basis to determine fixed export costs and to quantify their effects.

A hurdle model approach using firm-level data can separately identify how the existence of information costs and compliance costs may affect not only the extensive margin (i.e., the number of exporting firms), but also the intensive margin (i.e., the trade volume per exporting firm) of exporting firms. With industry-level data, the heterogeneity of exporting firms is represented by defining each SITC4-digit group in the form of "a representative firm".[3] Empirical results capture the estimated effects of information and compliance costs on decisions about industry export to differentiated markets.

The results show that the effects of fixed export costs are twofold. First, information costs are crucial in determining whether or not to export. Such costs decrease the probability of export by about five to six percentage points in the first stage. Once the export decision is made, firms tend to compensate information costs by exporting more in the second stage. On

[3] See Section 2.2.3 for discussions about how to fit industry level data to the firm level decision model.

the other hand, paying compliance costs decreases the probability of export by about 36 percentage points. Compliance costs are more prohibitive in the subsequent export value decision.

2.2. Empirical Framework

2.2.1 Firm Level Decision Process in the Presence of Information and Compliance Costs

When information costs are assumed to be zero and compliance costs are certain and known, the individual firm decision to enter a foreign market can be depicted as a simple two-step process (Figure 2.1.1). In time period t, a firm decides "whether or not to export" in the first stage t_1 (defined as the first sub period of t) given fixed export costs. In the second stage t_2 (defined as the second sub period of t) a firm decides how much to export. When collecting information is costly and compliance costs are known only after a firm collects the information, the export decision of an individual firm is still a two-step decision but with some modification in the first-stage t_1 as assumptions that fixed export costs are certain and known are relaxed. The decision of whether or not to export at time period t is now contingent on whether a firm must pay information costs and/or compliance costs in the first stage t_1 of time period t (Figure 2.1.2).

2.2.2 Econometric Model

Several approaches in the literature are used to model export decision and bilateral trade between countries. The most straightforward method to model individual firm export decision is the probit model (e.g., Roberts and Tybout, 1997; Bernard and Jenson, 2001). A log-linearized gravity equation is often used to predict the volume of trade for an individual exporting firm or the bilateral trade flow between countries. However, there is increasing evidence showing that zero trade flows are commonly observed in international trade. Haveman and Hummels (2004) find that nearly 1/3 of the bilateral trade matrix is empty. Helpmen et al. (2008) find that about half of the country pairs in their sample do not trade with each other. The problem is expected to be more serious at disaggregated firm level data when exporting zero volume is often an optimal choice. Under the log-linearized specification, taking logarithm effectively drops zero observations from the sample and is likely to produce biased estimates by getting rid of useful information. A commonly used empirical approximation is to add a small positive number (such as 0.0001) to all zero trade flows. This is sensible to see how including or excluding zeros make much of difference empirically, but has no theoretical basis.

A common econometric approach for dealing with corner solution (when some trade flows are zeros piling up at corner while others are strictly positive values) is the Tobit model (Tobin, 1958). Tobit model has been commonly applied in dealing with bilateral zero trade flows (e.g., Eaton and Tamura (1994)). [4]

In most recent empirical literature, in order to be consistent with firm heterogeneity theory (Melitz, 2003; Helpman et al., 2008), a two-step decision is often modeled in a Heckman procedure to account for bias that productive firms are self-selected into export market.

---

[4] Tobit model is not feasible in our case because the normality assumption required in the tobit model is not valid due to extremely large trade values observed in the data set.

In this paper, a lognormal hurdle model following Wooldridge (2010) is used to model the two-stage decision of firm participation in a foreign market.[5] The econometric model is parameterized as follows.

Let y be the export value chosen by an individual firm, which is a compound function of a binary participation decision variable s, and the continuous choice of a nonnegative export value y^*.

$$y = s \cdot y^* \tag{2.1}$$

When a firm decides to export to a specific foreign destination ($s = 1$), a nonnegative export value $y^* = y$ is observed. On the other hand, when a firm decides not to export ($s = 0$), then $y = 0$ and y^* is not observed.

The first-stage binary decision variable s is assumed to follow a probit model.

$$s = \begin{Bmatrix} 1 & \text{if a firm overcomes the hurdle of fixed export costs} \\ 0 & \text{otherwise} \end{Bmatrix} \tag{2.2.1}$$

Let s^* be the latent variable indicating unobserved firm ability to overcome market-specific fixed export costs which prevent a firm from participating in a foreign market. An individual firm will export only if it is able to overcome the hurdle of fixed export costs determined by a vector of attributes $\boldsymbol{x_1}$ such as firm characteristics including productivity

[5] The econometric model is directly formulated in a lognormal hurdle model because the trade values jump to large numbers after zeros at corner which fits better into a lognormal distribution than a truncated normal distribution. For details, see Section 17.6.2 in Wooldridge (2010).

and profit margins, compliance requirements in importing countries, etc. The observed attributes $\boldsymbol{x}_1$ is independent of the error term v, i.e., $cov(\boldsymbol{x}_1, v) = 0$.

$$s^* = \boldsymbol{x}_1\boldsymbol{\gamma} + v \qquad v \sim Normal\ (0,\ \ 1) \tag{2.2.2}$$

$$s = \begin{Bmatrix} 1 & s^* > 0 \\ 0 & s^* \leq 0 \end{Bmatrix} \tag{2.2.3}$$

$$P(s = 1|\boldsymbol{x}_1) = E(s|\boldsymbol{x}_1) = \Phi(\boldsymbol{x}_1\boldsymbol{\Upsilon}). \tag{2.2.4}$$

The continuous export value of y^* is then chosen based on a vector of attributes $\boldsymbol{x}_2$ such as firm characteristics, trade barriers, etc. Further, y^* is assumed to have a lognormal distribution (see footnote 6),

$$y^* = exp(\boldsymbol{x}_2\boldsymbol{\beta} + u) > 0 \quad u \sim Normal\ (0,\ \sigma^2) \tag{2.3.1}$$

$$y = \begin{Bmatrix} y^* > 0 & if \quad s = 1 \\ 0 & otherwise \end{Bmatrix}. \tag{2.3.2}$$

Or,

$$y = 1[\boldsymbol{x}_1\boldsymbol{\gamma} + v]exp(\boldsymbol{x}_2\boldsymbol{\beta} + u) > 0. \tag{2.4}$$

Let ρ be the correlation between v and u. Under the assumption that the binary export decision is independent of the export value decision conditioning on observed variables $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, $\rho = 0$. As $y = y^*$ when $y > 0$, the expression of the conditional density of y when $\rho = 0$ is,

$$f(y|\boldsymbol{x}_2, y > 0) = \phi\,[(log(y) - \boldsymbol{x}_2\boldsymbol{\beta})/\sigma]/(\sigma y)\,,\ \ y > 0.$$

(2.5)

The unconditional density of y given $\boldsymbol{x}_1,\ \boldsymbol{x}_2$ is straightforward by multiplying

$$P(y > 0|\boldsymbol{x}_1) = \Phi(\boldsymbol{x}_1\boldsymbol{\Upsilon}),$$

$$f\big(y|\boldsymbol{x}_1,\ \boldsymbol{x}_2\big) = [1 - \Phi(\boldsymbol{x}_1\boldsymbol{\gamma})]^{1[y=0]}$$

$$\{\Phi(\boldsymbol{x}_1\boldsymbol{\gamma})\phi\,[(log(y) - \boldsymbol{x}_2\boldsymbol{\beta})/\sigma]/(\sigma y)\}^{1[y>0]}.$$

(2.6)

For a random firm i, the associated log-likelihood function to be estimated is

$$l_i(\boldsymbol{\theta}) = 1[y_i = 0]\log[1 - \Phi(\boldsymbol{x}_i\boldsymbol{\gamma})] + 1[y_i > 0]\log[\Phi(\boldsymbol{x}_i\boldsymbol{\gamma})]$$

$$+1[y_i > 0]\{\log\{\phi[(\log(y_i) - \boldsymbol{x}_i\boldsymbol{\beta})/\sigma]\} - \log(\sigma) -$$

$$\log(y_i)\}. \quad (2.7)$$

The conditional and unconditional expectations of y then can be derived,

$$E(y|\boldsymbol{x}_2, y > 0) = E(y^*|\boldsymbol{x}_2, s = 1) = exp(\boldsymbol{x}_2\boldsymbol{\beta} + \sigma^2/2).$$

(2.8)

$$E(y|\boldsymbol{x}_1, \boldsymbol{x}_2) = \Phi(\boldsymbol{x}_1\boldsymbol{\Upsilon})exp(\boldsymbol{x}_2\boldsymbol{\beta} + \sigma^2/2). \quad (2.9)$$

The lognormal hurdle model assumes the binary choice decision (s) of whether or not to export is independent of the export value decision as long as the attributes which determine these two decisions can be controlled through observable firm characteristics. It is however possible that some unobserved factors that affect an individual firm's export decision (s) are still affecting the export value decision (y^*). Hence the independence assumption is relaxed and an exponential type II Tobit (ET2T) model is developed to allow conditional correlation between s and y^*.[6] This implies the correlation ρ between v and u is not zero and the variance-covariance matrix for v and u is $\begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}$. In order to identify ρ, an exclusive restriction is needed so that the covariates determining the export decision strictly contain those affecting the export value decision. In other words, $\boldsymbol{x}_2$ is a strict subset of $\boldsymbol{x}_1$. When ρ is non-zero, the conditional density of y given $\boldsymbol{x}_1$ can be derived,

$$f(y|\boldsymbol{x}_1, y > 0) = \Phi\big([\boldsymbol{x}_1\boldsymbol{\gamma} + (\rho/\sigma)(y - \boldsymbol{x}_2\boldsymbol{\beta})](1-\rho^2)^{-1/2}\big)$$

$$\phi\,[(log(y) - \boldsymbol{x}_2\boldsymbol{\beta})/\sigma]/(\sigma y). \qquad (2.10)$$

The unconditional density of given $\boldsymbol{x}_{1,}\ \ \boldsymbol{x}_2$ is

$$f\big(y|\boldsymbol{x}_{1,}\ \boldsymbol{x}_2\big) = [1 - \Phi(\boldsymbol{x}_1\boldsymbol{\gamma})]^{1[y=0]}$$

$$\big\{\Phi\big([\boldsymbol{x}_1\boldsymbol{\gamma} + (\rho/\sigma)(y - \boldsymbol{x}_2\boldsymbol{\beta})](1 -$$

$$\rho^2)^{-1/2}\big)\phi\,[(log(y) - \boldsymbol{x}_2\boldsymbol{\beta})/\sigma]/(\sigma y)\big\}^{1[y>0]}.$$

[6] See details in Section 17.6.2 in Wooldridge (2010).

(2.11)

and the associated log-likelihood function to be estimated is,

$$l_i(\boldsymbol{\theta}) = 1[y_i = 0]\log[1 - \Phi(\boldsymbol{x}_i\boldsymbol{\gamma})]$$

$$+1[y_i > 0]\left\{\log\left[\Phi([\boldsymbol{x}_i\boldsymbol{\gamma} + (\rho/\sigma)(\log(y_i) - \boldsymbol{x}_i\boldsymbol{\beta})])(1-\rho^2)^{-1/2}\right]\right.$$

$$\left. + \log\{\phi[(\log(y_i) - \boldsymbol{x}_i\boldsymbol{\beta})/\sigma]\} - \log(\sigma) - \log(y_i)\right\}.$$

(2.12)[7]

2.2.3 The Data

The major data used in the hurdle model are panel bilateral trade flow data set, selected from "World Trade Flows: 1962-2000" compiled by Feenstra and Lipsey.[8] The most recent 10 years (1991-2000) were selected as the analysis period and to be consistent with data information on regulation variables and trade barrier variables used in the empirical model.[9] Trade flows are recorded at the 4-digit SITC level using information from the importing countries' data sources wherever they are available for each importing-exporting country pair.[10] Under each particular 4-digit SITC code, trade values are summed to obtain the total trade flow corresponding to a specific SITC 4-digit product for each country pair in

---

[7] For detailed derivation of Eq. (2.10), (2.11) and (2.12), please see Chapter 17 in Wooldridge (2010).

[8] Available for download from http://cid.econ.ucdavis.edu/.

[9] Data on regulation variables are only available for 1999. Historical data for trade variables were not available and data for 2009 is used.

[10] Exporting country's information is used instead only when the importing country's report is not available when the data are compiled.

each year.[11] The 111A and 111X 4-digit categories for all country pairs and all years are deleted to construct a consistent SITC4-digit panel. [12]

Regulation variables measure how costly it might be for a domestic firm to meet all legal requirements before it can enter an industry or operate a new business (Djankov et al. (2002)). Regulations imposed by an importing country affect not only domestic firms but also foreign exporting firms seeking to enter the market. These regulations affect fixed export costs of entry but not variable costs of repeated export. In the empirical model, regulation variables include (a) the number of legal procedures, (b) number of days, and (c) relative official costs (as a percentage of GDP per capita) required for a new exporting firm to legally operate a business in the importing country. These three variables satisfy the exclusive restriction and can be used to measure the restrictiveness of entry to an importing country and to predict the likelihood of export for a foreign exporting firm in the first stage.

On the other hand, trade barriers imposed by an importing country are applied to foreign exporting firms only, which affect not only fixed export costs, but also variable costs of foreign exporting firms. Trade facilitation indices include a Logistics Performance Index (a scale number of 1-5 with 1 indicating the worst performance), a Burden of Customs Procedures (a scale number of 1-7 with 1 indicating the worst, i.e., the largest burden), a Lead Time measure (number of days delayed in the importing countries) and a Document measure (number of documents required by the importing countries in order to export) in importing

---

[11] The data reports three forms of trade value: (1) value (in thousand US dollars), (2) value with unit of number and (3) value with unit of weight.

[12] The 111A and 111X 4-digit categories are artificially created by Feenstra and Lipsey to capture missing or miscellaneous 4-digit trade flows for certain years so that the aggregation of all 4-digit SITC codes would equal to the value of a higher 3-digit SITC code. With no detailed information on countries or products included in 111A or 111X, the categories are inconsistent between years.

countries in year 2009 (Word Development Indicators, 2011). A measure of tariff rate of importing countries is drawn from Kee et al. (2009). This overall tariff rate measure is a weighted sum of all tariff lines over all goods estimated Kee et al. (2009) using tariff data between 2000 and 2004.

Other variables are adopted from Helpman et al. (2008), based on factors commonly used in gravity models to explain bilateral trade flows.[13] These variables include country characteristics (i.e., GDP per capita for importing and exporting countries), geographic variables (i.e., distance between the importing-exporting country pairs, whether importing and exporting countries share a common border, whether the exporting country is land locked or an island), institutional variables (i.e., whether the country pair shares the same legal system, shares the same colonial origin, whether both countries are members of WTO) and cultural variables (i.e., whether the country pair has a similarity in religion composition, speak the same primary language).

The compiled data set includes more than 13 million observations representing ten years of bilateral trade flow data between 117 importing countries and 117 exporting countries. To focus the analysis on top importing markets with large import demands, the model incorporates trade flows between the 117 exporting countries (Table 2.2) and only the 30 largest importing markets (Table 2.3).

2.2.4 Theoretical Assumptions and Restrictions

Several theoretical assumptions and restrictions are made to use the aggregate industry-level SITC4-digit data in a firm-level export decision model. First, a representative firm assumption is applied. Firms are assumed to be homogeneous within each SITC4-digit

[13] Data set available at http://scholar.harvard.edu/melitz/publications.

group and make the same export decisions when facing country-specific fixed information costs and compliance costs. The STIC4-digit sector within a country is thus representative of one firm. As there are 900 SITC4-digit codes in the data set, there are 900 representative firms for 900 different types of industries.

Second, the representative firm is assumed to only produce product(s) within the SITC4-digit group it represents. For example, a textile firm may produce and export both curtains and bedspreads, but these two products belong to the same SITC 4-digit code. On the other hand, if a machine manufacturing firm produces and exports both milling machines and other food processing machineries that belong to two different SITC 4-digit codes, these two product lines are assumed to be operated independently and thus are treated as if they were two different firms. This ensures the validity of aggregating trade flow within each SITC4-digit group.

Third, the follower assumptions are imposed for information costs and compliance cost to distinguish information costs from compliance costs as they cannot be observed in the data set. Information updates are costless as long as a firm remains active in the export destination, but information becomes obsolete if the firm exits the market for two consecutive years. Therefore, as long as a firm is exporting to a specific destination in at least one of the last two periods, no information costs are required to reenter the same market in period t. Information costs are relatively low cost to update as long as a firm remains actively participating in a specific foreign market.[14] Once information about a particular market is collected it is known by the firm.

[14] This research treats an exporting firm as active exporting firm in a foreign market if a firm continues to export to the same market in consecutive years or exits the market for only one year due to temporary shocks but reenters the market in the following year.

Once a firm exits a specific foreign market in period $t - 1$, it must re-pay compliance costs upon reentry into the market in period t. In reality some compliance costs are variable and must be incurred in each period as long as export continues. For example, after establishment of a processing line or acquiring certification for testing, packaging or labeling as part of the fixed export costs, compliance costs may vary with output as well such as the labor and materials for testing, labeling or packaging carried out in each period. Or, if investing a new processing line or obtaining a certain certificate for export requires a large expenditure, a firm may pay a proportion of compliance costs in each period as principle and interest in each future period on an initial loan.

The model assumes fixed export costs, including information costs, compliance costs, set-up of local distribution channels, etc., are market dependent. Previous or current experiences of exporting to a foreign market j neither reduces the information costs nor the compliance costs for an individual firm to enter another foreign market i in period t. Research regarding foreign demand, regulations, and standards in a potential foreign market must be conducted prior to entry into a new foreign market. Product is subject to modifications in order to meet the demands of individual markets as well as to comply with the foreign standards. Empirical results show that experience in one market is not relevant or does not increase the probability of exporting to another market (Blanes-Cristóbal, 2008).

Following these assumptions, the information and compliance costs faced by an individual firm to export to a specific foreign market in period t is summarized in Table 2.1.

Table 2.1: Costs of Export Decision at Period t Contingent on Period $t-1$ and $t-2$ Status

Group	……	$t-2$	$t-1$	t	……
I		export	export	no information cost, no compliance costs	
II		not export	export	no information cost, no compliance costs	
III		export	not export	no information cost, compliance costs	
IV		not export	not export	information cost, compliance costs	

Based on the above four groups, two binary indicators are used to capture the effect of compliance costs as well as information costs. As firms in Group I and II pay neither information costs nor compliance costs while firms in Group IV must pay both information costs and compliance costs in order to export in period t, the combined effect of both information costs and compliance costs can be identified by generating a binary variable *both_Index*, which equals one if a firm belongs to Group I or II, and zero if a firm belongs to Group IV. Meanwhile, firms in Group III pay no information costs but only compliance costs while firms in Group I and II must pay neither information costs nor compliance costs in order to export in period t, the effect of compliance costs can be identified by generating a binary variable *Comp_Index*, which equals one if a firm belongs to Group III, and zero if a firm belongs to Group I or II. The difference between *both_Index* and *Comp_Index* captures the effect of information costs.

2.2.5. Specification of Estimation Equations for a Representative Firm

With these above mentioned theoretical restrictions, the aggregate bilateral SITC4-digit industry-level trade flow data will fit into a representative firm's two-stage decision equations.[1]

In the first stage a probit regression defines firm decision to enter individual export market.

$$\Pr(EXPORT_{fijt} = 1|\boldsymbol{x}_1) = \Phi(\gamma_0 + Index_{fijt}\gamma_1 + Gravity_{ijt}\,\boldsymbol{\gamma}_2 + Regulation_i\,\boldsymbol{\gamma}_3 + TB_i\,\boldsymbol{\gamma}_4) \tag{2.13}$$

Conditional on a firm export results from the first stage, an OLS regression for observations with positive export values only is run in the second stage to determine level of export.

$$\log TV_{fijt} = \beta_0 + Index_{fijt}\beta_1 + Gravity_{ijt}\,\boldsymbol{\beta}_2 + TB_i\boldsymbol{\beta}_3 + u_{fijt}. \tag{2.14}$$

The dependent variable *EXPORT* in the probit regression is a binary variable, corresponding to one if a representative individual firm f in country j exports to importing country i in year t, and equal to zero otherwise. The dependent variable $\log TV$ in the lognormal regression is the log of trade value of SITC4-digit product f of representative firm f in exporting country j to importing country i in year t if an export decision is made in the first stage (Equation 2.13). Based on the single-product firm assumption, the subscript f stands for both a representative exporting firm and a SITC4-digit product it exports

[1] The same estimating equations could be used for heterogeneous firms if firm-level data were available.

Index aims to pick up the effects of information costs and compliance costs on the likelihood of export in the first stage (Equation 2.13) and subsequent export value decision in the second stage (Equation 2.14). *Index* represents two variables: *Comp_Index* (measures the effect of compliance costs in the model) and *Both_Index* (measures the combined effect of both information and compliance costs in the model). By construction, *Comp_Index* is a binary variable to capture the effect of compliance costs by comparing firms paying no information costs but only compliance costs (Group III) with firms paying neither information cost nor compliance costs (Group I and II). *Both_Index* is a binary variable to capture the effect of both information costs and compliance costs by comparing firms paying both information cost and compliance costs (Group IV) with firms paying neither information cost nor compliance costs (Group I and II). The difference between the effects of *Comp_Index* and *Both_Index* obtained separately from the lognormal hurdle model captures the effect of information costs. As *Both_Index* contains the effect of both information costs and compliance costs, we expect to observe a larger negative impact of *Both_Index* than *Comp_Index* which contains the effect of compliance costs only.

Regulation is a vector of regulation cost variables indicating the restrictiveness of starting a new business for a domestic firm in importing country i. The three variables are number of legal procedures (*procedure*), number of legal days (*time*), and relative official costs, as a percentage of GDP per capita (*cost*). Since both foreign and domestic firms face these same regulations, they affect a foreign firm's decision on whether to enter the foreign market but not on how much to export. Therefore, the three variables in the *Regulation* vector are used as exclusion restrictions in exponential type II Tobit (ET2T) model to identify ρ and test whether the two decisions are correlated. Given that these regulatory hurdles are faced by an exporting firm in order to start a new business in an importing country, an increase in number of legal procedures or number of legal days, or an increase in relative official costs, is expected to decrease the likelihood of export for an exporting firm.

On the other hand, a foreign firm faces additional hurdles to export such as tariff and non-tariff barriers imposed by importing countries. Hence, the vector of trade barrier (*TB*) variables is included to control for these restrictions. Four trade facilitation indices, *logistics*, *customs*, *lead_days* and *documents* are used to measure the importing country's overall openness. Improvement in logistics and customs (i.e. a larger logistics and customs index because *logistics* and *customs* are scaled numbers with a large number indicating a better service) encourage export while delays in time and increases in the number of required documents discourage exports. The variable *OTRI_tariff* created by Kee et al. (2009), a measure of the overall trade restrictiveness of tariff, controls for tariff rate in each importing country. Increase in tariff rate should decrease both export probability and export value.

Gravity is a vector of country-pair specific variables includes: distance between importing and exporting countries (*log(GDP_EX)*), whether importing and exporting countries share a common border (*border*), the same legal system (*legal_system*), the same colonial origin (*colonial_tie*), whether an exporting country is land locked (*landlock_EX*) or an island (*island_EX*),[2] whether both countries are members of WTO (*WTO*), have a similarity in religion composition(*religion*), speak the same primary language (*language*). Additionally, importing country's GDP per capita (*log(GDP_IM)*) and exporting country's GDP per capita (*log(GDP_EX)*) are included as control for country size and market demand.

[2] In standard gravity model, island or landlock dummy is usually generated to indicate whether both importing and exporting countries are islands (landlocked). As only the 30-top importing countries are included in this analysis, island and landlock dummies are created for exporting countries only to indicate whether an exporting country is island or landlocked .

By standard recognition of gravity model, increase in distance between importing and exporting country, sharing a common border, the same legal system and the colonial origin of both countries, and both being members of WTO will increase the probability of trade as well as trade volume between importing and exporting countries. A landlocked exporting country exports less because of little access to ports while an islanded exporting country exports more because of abundant access to ports, limit resources and economic dependency. Cultural similarities in religion and language will also increase trade between importing and exporting countries. Country size has positive impact on trade between two countries.

2.3 Results and Discussion

Table 2.6.1 and 2.6.2 summarize the effect of compliance costs in the export decision equation and export value equation respectively. Similarly, Table 2.6.1 and 2.6.2 summarize the combined effect of both information costs and compliance costs. At a marginal effect basis, paying compliance costs in the first stage decreases the probability of export by about 35 to 36 percentage points while paying both information costs and compliance costs decreases the probability of export by about 41 percentage points depending on model specifications (Table 2.2). The difference between these two effects approximates the negative effect of information costs, i.e., paying information costs decreases the probability of export by about five to six percentage point.

The estimated marginal effect of compliance costs and combined effect are significantly larger in the export value equation, especially in the exponential type II Tobit (ET2T) model when the export decision and export value decision are jointly determined

(Row (iii) and (iv), Table 2.2). In the lognormal hurdle model (i.e., export decision and export value decision are assumed to be independent), once the hurdle is overcome, paying compliance costs reduces export value by about 170 percent while paying both information and compliance costs reduces export value by about 177 percent which implies paying information costs reduces export value by about seven percent. On the other hand, in the exponential type II Tobit (ET2T) model, paying compliance costs leads to a reduction of more than 210 percent in export value. Meanwhile paying both information and compliance costs leads to a reduction of more than 180 percent in export value, which is smaller than the effect of compliance costs. This result suggests that in a process where the export decision and export value decision are not independent ($\rho \neq 0$), the effect of information costs and compliance costs are not a simple additive relationship as assumed and observed in the lognormal hurdle model. Once information is obtained and the export decision is made, information costs may have positive impact in determining how much to export. In other words, comparing with a firm that pays compliance costs only, a firm that pays both information costs and compliance costs tends to export more in order to compensate the additional information costs it has to pay for export.

The compliance costs index (*Comp_Index*) and the combined index of both information costs and compliance costs (*Both_Index*) were generated based on theoretical assumptions regarding export status summarized in Table 2.1. By construction, the prediction power of these two indicators is limited to predicting the probability of export in the first stage. However, it is not surprising to observe such a significantly negative effect of information costs and compliance costs on export value. These marginal effects capture the difference between exporting firms (firms categorized in Group I and II export in period $t-1$) and non-exporting with zero export values (firms categorized in Group III and IV do not export in period $t-1$).

Table 2.2: Marginal Effects of Compliance Costs and Combined Information and Compliance Costs

Marginal Effect		Lognormal Hurdle Model (1)	Lognormal Hurdle Model (2)	ET2T Model (3)	ET2T Model (4)
Export Probability					
(i)	Compliance Costs Effect	-0.355** (0.001)	-0.351** (0.001)	-0.361** (0.001)	-0.358** (0.001)
(ii)	Combined Effect	-0.415** (0.000)	-0.412** (0.000)	-0.412** (0.000)	-0.410** (0.000)
Export Value					
(iii)	Compliance Costs Effect	-1.695** (0.006)	-1.686** (0.006)	-2.189** (0.006)[3]	-2.177** (0.006)[3]
(iv)	Combined Effect	-1.770** (0.005)	-1.762** (0.005)	-1.819** (0.004)[3]	-1.811** (0.004)[3]

Notes: 1. Marginal effect is the partial effect averaged across sample.
2. Standard errors are reported in parenthesis. **Significant at 1%.
3. To be comparable with the Lognormal Hurdle Model, zero trade values are excluded when calculating marginal effects.

As the independence assumption of the export decision and export value decision is strongly rejected at χ^2=32706.27 for compliance costs effect (Column 3 in Table 2.6.2), and at χ^2=45006.84 for combined effect of both information costs and compliance costs (Column 3 in Table 2.7.2) under the baseline specifications, discussion of results on other variables are focused the exponential type II tobit (ET2T) model.

For the three regulation variables, the results indicate a small negative effect of *time* and *cost* as expected but not *procedure* in the export decision equation. An increase in one hundred percent of relative official cost (as a percentage of GDP per capita) of starting up a new business in an importing country reduces the probability of export by less than one percentage point. An increase of ten days to start a new business in importing country reduces the probability of export by 0.2 percentage point. On the other hand, an additional procedure required in an importing country for a new business is found to increase the probability of exporting by 0.01 percentage point. As noted earlier, it was expected that the three regulation variables would have the same inverse relationship with the likelihood of export as they all measure the entry costs faced by an exporting firm. One possible reason for this unexpected positive effect of *procedure* on export value is that changes in the number of procedures in importing countries are not fully captured because the three regulation variables are only available for year 1999 while the analysis period is from 1991 to 2000.

For trade barrier variables, an improvement of logistics in an importing country encourages foreign firms to export and export more while improvement of customs procedure in an importing country is unexpectedly found to discourage exports. Time delay in execution of foreign firm's exporting procedure decreases export value but increases the probability of export. An increase in the number of required documents regarding exporting decreases both the probability of export and export value. An increase in tariff rate (*OTRI_tariff*) leads to less export and reduced export value. Comparing with information costs and compliance costs, the negative effect due to tariff rate is minimal.

Gravity variables have a significant effect in the export decision equation and in the export value equation. In particular, an increase in the distances between the country pair reduces exports. Both importing and exporting country size matters as both the probability of exporting and export value increase as country size increases. In addition, the effect of country size is asymmetric in such a way that the size of an exporting country has a larger impact on both export decision and value decision. Meanwhile, exporting and importing countries sharing a border, a common language or a same legal system will increase exports between the countries. Exports increase for an island country, but decrease for a landlocked country, keeping other conditions constant.

It is notable that the estimated coefficients on the set of gravity variables are not all consistent with the standardized results which are obtained from a single gravity equation frequently utilized to either predict the probability of export or the bilateral traded flow. In particular, both importing and exporting countries being members of WTO, or having a colonial relationship, or sharing similarities in religions does not always increase the individual firm's export probability and export value. The discrepancy with standardized results estimated from aggregated country level data could be attributable to the fact that the model is at firm level but an aggregated industry level data are used to estimate the firm level export decision.

The inconsistency of predicted effects for some of the regulation variables, trade barrier variables and gravity variables might be caused by the limited scope of this study.[3] An aggregated data are used to approximate the firm level export decision. In addition, only the top 30 importing countries are considered and the characteristics of exporting countries are not controlled in this study. [4] In addition, data on regulation, trade barriers as well as country size has no time variation. Changes in regulatory environment, trade policies and economic development in the 30 importing countries over the analysis period are not captured.[5]

---

[3] The scope limitation does not affect the estimated effects of compliance costs and information costs. Results are not significantly altered by experiments to drop different gravity variables, year fixed effect and SITC2-digit fixed effect.

[4] Exporting country fixed effect is not feasible due to collinearity caused by excessive dummies in the regression.

[5] Interacting with year dummy may help resolve this problem. Unfortunately, once interaction terms added, log likelihood function is not concave and does not converge.

2.4 Conclusion

When compliance costs are unknown until an individual firm pays information costs prior to exporting, the export decision is no longer as simple as if it were facing fixed export costs that are certain. Export decisions depend on whether information costs and/or compliance costs are required in order to export in the current period. This research investigates the effect of information costs and compliance costs on a firm's decisions to export and how much to export through a hurdle model approach. By fitting a panel SITC 4-digit bilateral trade flow data into an empirical representative firm-level export decision model, this research identified and decomposed the different effects of information costs and compliance costs in firms' export decision and subsequent export value decision. The effects of information costs and compliance costs are two-fold. Information costs are crucial in determining whether or not to export to a specific foreign market. It decreases the probability of export by about five to six percentage points in the first stage. Once the export decision is made, firms paying information costs tend to export more in the second stage to cover such costs. On the other hand, paying compliance costs decreases the probability of export by about 36 percentage points. Compliance costs are more prohibitive in the subsequent export value decision.

Due to data limitations, within-industry heterogeneity of exporting firms was ignored by imposing some theoretical restrictions in order to fit the aggregate data to the firm-level regression model. The true effect of information costs and compliance costs on heterogeneous firm's export and export value decisions would be larger if ignoring within-industry heterogeneity of exporting firms causes any down-ward bias in the estimation. Whenever a firm-level panel is available, restrictions imposed in models in this study are redundant. The logic as well as the estimating procedures can be duplicated to quantify the effect of information costs as well as compliance costs on the export and export value decisions of heterogeneous firms.

APPENDIX

Table 2.3: List of 117 Exporting Countries and Areas

Albania	Ghana	Norway
Algeria	Greece	Oman
Angola	Guatemala	Pakistan
Argentina	Guinea	Panama
Australia	Haiti	Papua N.Guinea
Austria	Honduras	Paraguay
Bangladesh	Hong Kong, China	Peru
Belgium–Lux.	Hungary	Philippines
Benin	India	Poland
Bhutan	Indonesia	Portugal
Bolivia	Iran	Romania
Brazil	Ireland	Rwanda
Bulgaria	Israel	Saudi Arabia
Burkina Faso	Italy	Senegal
Burundi	Jamaica	Sierra Leone
Cambodia	Japan	Singapore
Cameroon	Jordan	Solomon Islds.
Canada	Kenya	South Africa
Central Africa Republic	Kiribati	Spain
Chad	Korea Rep. (South)	Sri Lanka
Chile	Kuwait	Sweden
China	Laos	Switzerland
Colombia	Lebanon	Syrn Arab Rp
Congo	Madagascar	Taiwan, China
Costa Rica	Malawi	Tanzania
Cote D’Ivoire	Malaysia	Thailand
Czech Republic (Fm. Czechoslovakia)	Maldives	Togo
Denmark	Mali	Tunisia
Dominican Rep.	Mauritania	Turkey
Ecuador	Mexico	Uganda
Egypt	Mongolia	United Kingdom
El Salvador	Morocco	United Arab Em.
Ethiopia	Mozambique	United States
Fiji	Nepal	Uruguay
Finland	Netherlands	Venezuela
Fm. USSR	New Zealand	Vietnam
Fm. Yugoslavia	Nicaragua	Yemen
France	Niger	Zambia
Germany	Nigeria	Zimbabwe

Table 2.4: List of Top 30 Importing Countries*

Australia (island)	Mexico
Austria	Norway
Belgium–Lux	Portugal
Brazil	Russia
Canada	Saudi Arabia
China	Singapore (island)
Czech Republic (landlocked)	Spain
France	Sweden
Germany	Switzerland (landlocked)
India	Thailand
Indonesia (island)	Turkey
Italy	United Kingdom (island)
Japan (island)	United Arab Em.
Korea Rep. (South)	United States
Malaysia	Vietnam

*Notes: Countries are listed in alphabetical order. The rank is based on World Trade Report 2012 by WTO (http://www.wto.org/english/res_e/booksp_e/anrep_e/world_trade_report12_e.pdf)

Table 2.5: Descriptive Statistics of Variables

Variable Name		Definition	No. of Observations	Mean	Std. Dev.	Min	Max
Complinace_Index		Binary, 1 if a firm doesn't pay information costs but does pay compliance costs in order to export to a specific destination, zero otherwise	2,012,587	0.121	0.326	0	1
Both_Index		Binary, 1 if a firm pays both information costs and compliance costs in order to export to a specific destination, zero otherwise	4,493,027	0.606	0.489	0	1
EXPORT		Binary, 1 if a firm exports to a specific destination, 0 otherwise	5,919,840	0.373	0.484	0	1
Log(TV)		Log of yearly trade value between a importing and exporting county pair at SITC4-digit level	2,208,910	14.071	1.821	6.908	24.261
Regulation	Procedure	No. of procedures required for a domestic start-up firm to legally operate a business in the importing country	5,919,840	8.039	3.355	2	17
	Time	No. of official days required for a domestic start-up firm to legally operate a business in the importing country	5,919,840	38.903	37.580	2	152
	Cost (%)	Official costs required for a domestic start-up firm to legally operate a business in the importing country (% of GDP per capita)	5,919,840	16.517	25.031	0.6	130.7
Trade Barrier	Logistics	Logistics Performance Index (a scale of 1-5 with 1 indicating the worst performance	5,285,150	3.636	0.376	2.61	4.11
	Customs	Burden of Customs Procedures (a scale number of 1-7 with 1 indicating the largest burden),	5,285,150	4.442	0.997	0.6	5.8
	Lead_Days	No. of days between initiation and execution of a exporting process in importing country	5,285,150	3.414	1.484	1	7.1
	Documents	No. of documents required by the importing countries in order to export	5,285,150	5.108	1.4870	2	13

Table 2. 5 (Cont'd)

	OTRI_tariff (%)	Overall trade restrictiveness index (tariff data only)	5,285,150	5.615	5.560	1.7	26.1
Gravity	Log(GDP_IM)	Log of GDP per capita in importing country (in 1999 US dollars)	5,919,840	9.401	1.227	6.174	10.677
	Log(GDP_EX)	Log of GDP per capita in exporting countries (in 1999 US dollars)	5,908,740	9.060	1.425	4.500	10.677
	Log(distance)	Log of the distance (in km) between importing and exporting country's capital	5,919,840	3.908	1.016	0.882	5.661
	Border	Binary, 1 if importing and exporting country shares a common border, 0 otherwise	5,919,840	0.069	0.253	0	1
	Island_EX	Binary, 1 if exporting country is islands, 0 otherwise	5,919,840	0.106	0.308	0	1
	Landlock_EX	Binary, 1 if exporting country has no direct access to sea, 0 otherwise	5,919,840	0.114	0.318	0	1
	Language	Binary, 1 if both importing and exporting country use the same language as official language	5,919,840	0.214	0.410	0	1
	Legal system	Binary, 1 if both importing and exporting country share the same legal origin	5,919,840	0.276	0.447	0	1
	Religion	(% Protestants in importing country* % Protestants in exporting country)+(% Catholics in importing country * %Catholics in exporting country) + (% Muslims in importing country* %Muslims in exporting country)	5,919,840	0.174	0.254	0	0.987
	Colonial tie	Binary, 1 if importing country ever colonized exporting country or vice versa, 0 otherwise.	5,919,840	0.044	0.206	0	1
	WTO	Binary, 1 if both exporting and importing country belong to WTO, 0 otherwise	5,919,840	0.739	0.439	0	1

Table 2.6.1: Compliance Costs Effect in Export Decision Equation

Explanatory Variable	Export Decision Equation Dependent Variable: *Export*							
	Lognormal Hurdle Model				ET2T Model (Heckman Method)			
	(1)		(2)		(3)		(4)	
Compliance Index	-1.609**	(0.004)	-1.601**	(0.003)	-1.618**	(0.004)	-1.611**	(0.004)
No. of procedures	0.004**	(0.001)	0.002*	(0.001)	0.008**	(0.000)	0.009**	(0.000)
Time	-0.001**	(0.000)	-0.001**	(0.000)	-0.001**	(0.000)	-0.001**	(0.000)
Cost (%)	0.001**	(0.000)	0.001**	(0.000)	-0.000**	(0.000)	-0.000	(0.000)
Logistics	0.204**	(0.010)	0.079**	(0.012)	0.331**	(0.010)	0.194 **	(0.012)
Customs	-0.075**	(0.002)	-0.042**	(0.003)	-0.087**	(0.002)	-0.070**	(0.002)
Lead Time	-0.027**	(0.001)	-0.003**	(0.001)	0.001	(0.001)	0.003^{+}	(0.002)
Documents	-0.042**	(0.001)	-0.007**	(0.002)	-0.071**	(0.001)	-0.063**	(0.002)
OTRI_tariff	-0.006**	(0.000)	0.020**	(0.001)	-0.006**	(0.000)	0.003**	(0.001)
Log(GDP_IM)	0.031**	(0. 003)	-0.044**	(0.005)	0.006*	(0.003)	0.007	(0.005)
Log(GDP_EX)	0. 114**	(0. 001)	0.047**	(0.004)	0.115**	(0.001)	0.105**	(0.005)
Log(distance)	-0. 019**	(0.002)	-0.031**	(0.002)	-0.020**	(0.002)	-0.023**	(0.002)
Border	0. 329**	(0.007)	0.328**	(0.007)	0.505**	(0.008)	0.505**	(0.008)
Island_EX	0.141**	(0.005)	0.136**	(0.005)	0.156 **	(0.005)	0.150**	(0.005)
Landlock_EX	-0.150**	(0.005)	-0.139**	(0.005)	-0.223**	(0.005)	-0.227**	(0.005)
Language	0. 018**	(0.004)	0.034**	(0.004)	0.044**	(0.004)	0.063**	(0.004)
Legal System	0.036**	(0.004)	0.027**	(0.004)	0.040**	(0.004)	0.032**	(0.004)
Religion	-0.024**	(0.007)	0.035**	(0.007)	-0.106**	(0.007)	-0.058**	(0.007)
Colonial tie	-0.041**	(0.008)	-0.016*	(0.008)	-0.042**	(0.009)	-0.040**	(0.009)
WTO	-0.027**	(0.004)	0.017**	(0.005)	-0.034**	(0.004)	-0.042**	(0.005)

Table 2.6.1(Cont'd)

Year fixed effects	Yes	Yes	Yes	Yes
SITC2-digit fixed effects	Yes	Yes	Yes	Yes
Importing country income group fixed effects[a,b]	-	Yes	-	Yes
Exporting country income group fixed effects[a,b]	-	Yes	-	Yes
Observations	1,783,220	1,783,220	1,783,221	1,783,221
Log likelihood	-713,346	-708,650	-3,416,187	-3,410,884

Notes: Robust standard errors (clustering by country pair SITC 4-digit level) are reported in parenthesis.

+Significant at 10%.

*Significant at 5%.

**Significant at 1%.

a. Importing and exporting country fixed effects are suppressed as the log likelihood function does not converge.

b. According to World Bank estimates of 1999 GNP per capita. Low income group: $755 or less; Lower middle income group: $756-2,995; Upper middle income group: $2,996-9,265; High income group: $9,266 or more.

Table 2.6.2: Compliance Costs Effect in Export Value Equation

Explanatory Variable	Export Value Equation Dependent Variable: log(TV)							
	Lognormal Hurdle Model				ET2T Model (Heckman Method)			
	(1)		(2)		(3)[c]		(4)[d]	
Index	-1.695**	(0.006)	-1.686**	(0.006)	-3.374**	(0.008)	-3.350**	(0.009)
No. of procedures	-	-	-	-	-	-	-	-
Time	-	-	-	-	-	-	-	-
Cost (%)	-	-	-	-	-	-	-	-
Logistics	0.384**	(0.022)	0.078**	(0.027)	0.409**	(0.022)	0.070**	(0.028)
Customs	-0.162**	(0.004)	-0.150**	(0.005)	-0.200**	(0.004)	-0.153**	(0.005)
Lead Time	-0.065**	(0.003)	-0.079**	(0.004)	-0.094**	(0.003)	-0.089**	(0.004)
Documents	-0.025**	(0.003)	-0.022**	(0.003)	-0.040*	(0.003)	-0.015**	(0.003)
OTRI_tariff	-0.005**	(0.001)	-0.002	(0.002)	-0.010**	(0.001)	0.012**	(0.002)
Log(GDP_IM)	0.043**	(0.007)	0.096**	(0.012)	0.080**	(0.007)	0.073**	(0.012)
Log(GDP_EX)	0.121**	(0.003)	0.248**	(0.011)	0.203**	(0.003)	0.262**	(0.011)
Log(distance)	-0.023**	(0.004)	-0.024**	(0.004)	-0.033**	(0.004)	-0.041**	(0.004)
Border	0.992**	(0.016)	0.996**	(0.017)	1.130**	(0.017)	1.133**	(0.017)
Island_EX	0.209*	(0.011)	0.222**	(0.011)	0.290**	(0.012)	0.304**	(0.011)
Landlock_EX	-0.405**	(0.011)	-0.442**	(0.011)	-0.481**	(0.011)	-0.504**	(0.011)
Language	0.095**	(0.009)	0.119**	(0.009)	0.103**	(0.010)	0.136**	(0.010)
Legal System	0.069**	(0.009)	0.065**	(0.009)	0.088**	(0.009)	0.077**	(0.009)
Religion	-0.266**	(0.015)	-0.232**	(0.015)	-0.268**	(0.016)	-0.213**	(0.016)
Colonial tie	-0.079**	(0.019)	-0.091**	(0.019)	-0.110**	(0.019)	-0.095**	(0.019)
WTO	-0.218**	(0.001)	-0.288**	(0.011)	-0.244**	(0.010)	-0.265**	(0.011)

Table 2.6.2 (cont'd)

Year fixed effects	Yes	Yes	Yes	Yes
SITC2-digit fixed effects	Yes	Yes	Yes	Yes
Importing country income group fixed effects[a, b]	-	-	Yes	Yes
Exporting country income group fixed effects[a, b]	-	-	Yes	Yes
Observations	1,419,403	1,743,634	1,783,221	1,783,221
Log likelihood	-28,823,138	-28,816,311	-3,416,187	-3,410,884

Notes: Robust standard errors (clustering by country pair SITC4-digit level) are reported in parenthesis.

+Significant at 10%.

*Significant at 5%.

**Significant at 1%.

a. Importing and exporting country fixed effects are suppressed as the log likelihood function does not converge.

b. According to World Bank estimates of 1999 GNP per capita. Low income group: $755 or less; Lower middle income group: $756-2,995; Upper middle income group: $2,996-9,265; High income group: $9,266 or more.

c. Independent assumption of two stages($\rho = 0$) is rejected withχ^2=30961.78.

d. Independent assumption of two stages($\rho = 0$) is rejected withχ^2=31144.09.

Table 2.7.1: Combined Information Costs and Compliance Costs Effect in Export Decision Equation

Explanatory Variables	Export Decision Equation Dependent Variable: *Export*							
	Lognormal Hurdle Model				ET2T Model (Heckman Method)			
	(1)		(2)		(3)		(4)	
Both Index	-2.502**	(0.002)	-2.494**	(0.002)	-2.480**	(0.002)	-2.473**	(0.003)
No. of procedures	-0.001**	(0.000)	-0.002**	(0.000)	0.005**	(0.005)	0.006**	(0.000)
Time	-0.001**	(0.000)	-0.002**	(0.000)	-0.001**	(0.000)	-0.001**	(0.000)
Cost (%)	0.001**	(0.000)	0.001**	(0.000)	-0.000**	(0.000)	-0.000**	(0.000)
Logistics	-0.012**	(0.005)	-0.064**	(0.006)	0.091**	(0.006)	-0.033**	(0.007)
Customs	-0.109**	(0.001)	-0.083**	(0.002)	-0.081**	(0.001)	-0.090**	(0.002)
Lead Time	-0.015**	(0.001)	0.003**	(0.001)	0.001	(0.001)	0.005**	(0.001)
Documents	-0.069**	(0.001)	-0.052**	(0.001)	-0.082*	(0.001)	-0.077**	(0.001)
OTRI_tariff	0.004**	(0.000)	0.022**	(0.000)	0.002**	(0.000)	0.009**	(0.000)
Log(GDP_IM)	-0.005**	(0.002)	-0.041**	(0.003)	-0.003	(0.002)	0.019**	(0.004)
Log(GDP_EX)	0.054**	(0.001)	-0.037**	(0.003)	0.066**	(0.001)	0.021**	(0.003)
Log(distance)	0.003**	(0.001)	-0.001**	(0.001)	-0.000	(0.001)	-0.003*	(0.001)
Border	0.068**	(0.004)	0.060**	(0.004)	0.242**	(0.005)	0.242**	(0.005)
Island_EX	0.096*	(0.003)	0.091**	(0.003)	0.119**	(0.004)	0.114**	(0.004)
Landlock_EX	-0.071**	(0.003)	-0.061**	(0.003)	-0.142**	(0.003)	-0.142**	(0.003)
Language	0.105**	(0.002)	0.108**	(0.002)	0.110**	(0.003)	0.121**	(0.003)
Legal System	0.063**	(0.002)	0.060**	(0.002)	0.065**	(0.003)	0.061**	(0.003)
Religion	0.006	(0.004)	0.042**	(0.004)	-0.062**	(0.005)	-0.032**	(0.005)
Colonial tie	-0.130**	(0.004)	-0.113**	(0.004)	-0.118**	(0.006)	-0.116**	(0.006)
WTO	0.043**	(0.002)	0.077**	(0.003)	0.019**	(0.003)	0.014**	(0.003)

Table 2.7.1 (cont'd)

Year fixed effects	Yes	Yes	Yes	Yes
SITC2-digit fixed effects	Yes	Yes	Yes	Yes
Importing country income group fixed effects[a, b]	-	Yes	-	Yes
Exporting country income group fixed effects[a, b]	-	Yes		Yes
Observations	3,978,791	3,978,791	4,007,318	4,007,318
Log likelihood	-1,220,891	-1,216,647	-4,083,178	-4,078,049

Notes: Robust standard errors (clustering by country pair SITC4-digit level) are reported in parenthesis.
+Significant at 10%.
*Significant at 5%.
**Significant at 1%.

a. Importing and exporting country fixed effects are suppressed as the log likelihood function does not converge.
b. According to World Bank estimates of 1999 GNP per capita. Low income group: $755 or less; Lower middle income group: $756-2,995; Upper middle income group: $2,996-9,265; High income group: $9,266 or more.

Table 2.7.2: Combined Information Costs and Compliance Costs Effect in Export Value Equation

Explanatory Variable	Export Value Equation Dependent Variable: log(TV)							
	Lognormal Hurdle Model				ET2T Model (Heckman Method)			
	(1)		(2)		(3)[c]		(4)[d]	
Both Index	-1.770**	(0.005)	-1.763**	(0.004)	-4.973**	(0.009)	-4.950**	(0.009)
No. of procedures	-	-	-	-	-	-	-	-
Time	-	-	-	-	-	-	-	-
Cost (%)	-	-	-	-	-	-	-	-
Logistics	0.346**	(0.020)	0.050**	(0.025)	0.172**	(0.017)	-0.159**	(0.020)
Customs	-0.154**	(0.004)	-0.143**	(0.004)	-0.217**	(0.004)	-0.172**	(0.004)
Lead Time	-0.065**	(0.003)	-0.080**	(0.003)	-0.094**	(0.003)	-0.089**	(0.003)
Documents	-0.020**	(0.003)	-0.017**	(0.003)	-0.057**	(0.002)	-0.037**	(0.003)
OTRI_tariff	-0.005**	(0.001)	-0.001	(0.001)	-0.000	(0.001)	0.020**	(0.001)
Log(GDP_IM)	0.039**	(0.006)	0.084**	(0.011)	0.066**	(0.005)	0.072**	(0.010)
Log(GDP_EX)	0.110**	(0.003)	0.227**	(0.010)	0.145**	(0.002)	0.145**	(0.009)
Log(distance)	-0.021**	(0.004)	-0.020**	(0.004)	-0.010**	(0.004)	-0.018**	(0.004)
Border	0.971**	(0.016)	0.975**	(0.016)	0.877**	(0.015)	0.882**	(0.015)
Island_EX	0.203*	(0.011)	0.216**	(0.011)	0.255**	(0.010)	0.266**	(0.010)
Landlock_EX	-0.374**	(0.010)	-0.408**	(0.010)	-0.370**	(0.009)	-0.385**	(0.009)
Language	0.087**	(0.009)	0.112**	(0.009)	0.178**	(0.008)	0.206**	(0.008)
Legal System	0.065**	(0.008)	0.062**	(0.008)	0.124**	(0.007)	0.117**	(0.007)
Religion	-0.246**	(0.013)	-0.211**	(0.014)	-0.202**	(0.012)	-0.164**	(0.013)
Colonial tie	-0.082**	(0.008)	-0.094**	(0.017)	-0.222**	(0.015)	-0.209**	(0.015)
WTO	-0.201**	(0.009)	-0.266**	(0.009)	-0.155**	(0.008)	-0.173**	(0.008)

Table 2.7.2 (cont'd)

Year fixed effects	Yes	Yes	Yes	Yes
SITC2-digit fixed effects	Yes	Yes	Yes	Yes
Importing country income group fixed effects	-	Yes	-	Yes
Exporting country income group fixed effect	-	Yes	-	Yes
Observations	1,533,769	1,533,769	4,007,318	4,007,318
Log likelihood	-29,518,233	-29,511,899	-4,083,178	-4,078,049

Notes: Robust standard errors (clustering by country pair SITC4-digit level) are reported in parenthesis.
+Significant at 10%.
*Significant at 5%.
**Significant at 1%.

a. Importing and exporting country fixed effects are suppressed as the log likelihood function does not converge.
b. According to World Bank estimates of 1999 GNP per capita. Low income group: $755 or less; Lower middle income group: $756-2,995; Upper middle income group: $2,996-9,265; High income group: $9,266 or more.
c. Independent assumption of two stages($\rho = 0$) is rejected with χ^2 =45006.84.
d. Independent assumption of two stages($\rho = 0$) is rejected with χ^2=45368.32.

Figure 2.1.1: Timeline for firm export decisions when fixed export costs are certain

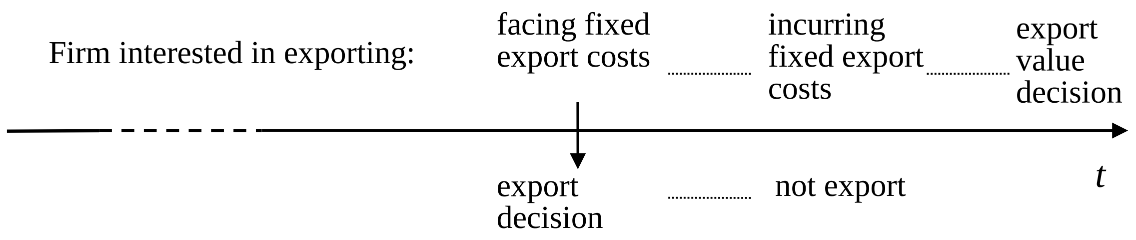
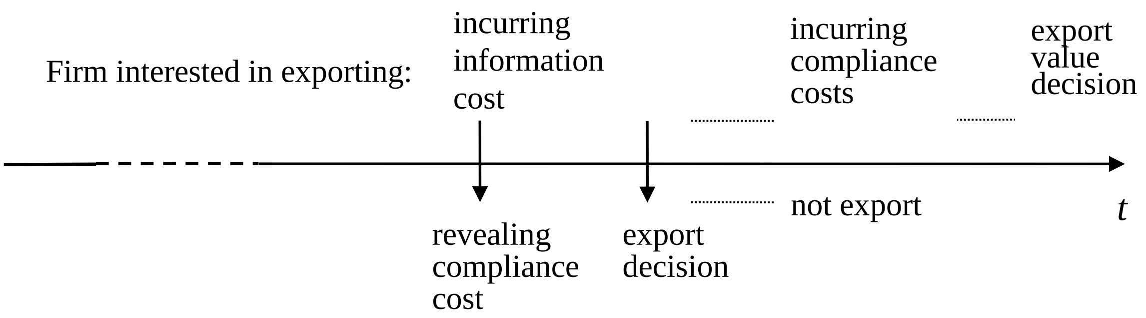


Figure 2.1.2: Timeline for firm export decisions when there is uncertainty in fixed export costs, such as compliance costs



REFERENCES

REFERENCES

Bernard, A. B. and J. B. Jensen (2004). Why some firms export. *The Review of Economics and Statistics 86* (2), 561–69.

Blanes-Cristóbal, J. V., M. Dovis, J. Milgram-Baleix, and A. I. Moro-Egido (2008). Do sunk exporting costs differ among markets? evidence from spanish manufacturing firms. *Economics Letters 101* (2), 110 – 12.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica 39* (5), 829–44.

Das, S., M. J. Roberts, and J. R. Tybout (2007). Market entry costs, producer heterogeneity, and export dynamics. *Econometrica 75* (3), 837–73.

Eaton, J. and A. Tamura (1994). Bilateralism and regionalism in Japanese and US trade and direct foreign investment patterns. *Journal of the Japanese and international economies 8* (1), 478–510.

Feenstra, R. C., R. E. Lipsey, H. Deng, A. C. Ma, and H. Mo (2005). World trade flows: 1962-2000. Working Paper 11040, National Bureau of Economic Research.

Haveman, J., & Hummels, D. (2004). Alternative Hypotheses and the Volume of Trade: The Gravity Equation and the Extent of Specialization. *Canadian Journal of Economics* 37 (1), 199-218.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics 123* (2), 441–87.

Kee, L. H., A. Nicita, and M. Olarreaga (2009). Estimating trade restrictiveness indices. *The Economic Journal 119* (534), 172–99.

Maskus, K. E., T. Otsuki and J. S. Wilson. "An empirical framework for analyzing technical regulations and trade," in *Quantifying the Impact of Technical Barriers to Trade,* ed. Keith E. Maskus and Jone S. Wilson (Ann Arbor : University of Michigan Press, 2001), 29.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71* (6), 1695–725.

Morales, E., G. Sheu, and A. Zahler (2011). Gravity and extended gravity: estimating a structural model of export entry. MPRA Paper 30311, University Library of Munich, Germany.

Roberts, M. J. and J. R. Tybout (1997). The decision to export in Colombia: An empirical model of entry with sunk costs. *American Economic Review 87* (4), 545–64.

Thornsbury, S., D. Roberts, and D. Orden. Measurement and political economy of disputed technical regulations. *Journal of agricultural and applied economics 36* (3).

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica 26* (1), 24–36.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data, Volume 1. The MIT Press.

CHAPTER 3: A STRUCTURAL ESTIMATION OF THE EMPLOYMENT EFFECTS OF OFFSHORING IN THE U.S. LABOR MARKET

3.1 Introduction

The debate over offshoring has intensified in the United States as offshoring has spread from the jobs of blue-collar workers in the manufacturing sector to those of white-collar workers in service sectors. The service sector comprises about 80 percent (U.S. Department of Commerce) of the U.S. employment and most white-collar workers are employed in the service sector. U.S. workers in all sectors have become more concerned about the security of their jobs due to increased offshoring activities as the global economy has continued to integrate. Given changes in technology (the internet), a well-educated radiologist and a low-skilled auto assembly line worker could both be susceptible to offshoring. These concerns are well reflected in results from Princeton University's telephone survey conducted in summer 2008.[1] Survey results indicate occupational offshorability reported by individual survey respondents are much higher than those predicted by economists.

The increase in offshoring along with a persistently high unemployment rate in recent years, has heightened policymaker concerns and has been the subject of increased economic research on the short- and long-run labor market implications of offshoring and in particular, the potential for U.S. job loss. The actual impact of offshoring is multi-dimensional and difficult to quantify. Existing empirical estimates (Bardhan and Kroll, 2003; Blinder, 2007; Blinder, 2009) provide a wide range of estimates for offshorable jobs in the U.S. labor market, varying from 11 to 47 percent. With relatively little theoretical

[1] For details, see Blinder and Krueger (2009).

guidance, the wide range in early empirical estimates provided limited information to policymakers facing tensions from a high national unemployment rate exceeding nine percent.

Under such circumstances, an economic theory of offshoring has been exposited by Grossman and Rossi-Hansberg (2008). In their parsimonious framework, job tasks are defined as either low-skilled or high-skilled. Using comparative static analysis, they analyze the synergic action of *productivity effect*, *relative-price effect* and *labor supply effect* of offshoring on these two groups due to a change of offshoring costs. Their results show that offshoring might lead to wage gains for both low-skilled and high-skilled workers and create a win-win situation for all types of workers, but not necessarily reward one player by harming the others as stated in the traditional Stolper-Samuelson results. Motivated by these results, several papers empirically tested the effect of offshoring in the United States (Harrison and McMillan, 2010; Ebenstein et al., 2013; Crinò, 2010b) and in European countries (Goos et al., 2010; Crinò, 2010a; Criscuolo and Garicano, 2010).

Harrison and McMilan (2010) estimated a reduction of four million jobs in U.S. manufacturing employment due to offshoring over the period of 1982 to 1999. Ebenstein et al. (2013) found that the impact of offshoring on U.S. worker's wages has been underestimated by previous studies because offshoring has driven workers from high-wage manufacturing jobs to low-wage service jobs. In addition, workers performing routine tasks are most affected by offshoring and experience larger wage decline. On the other hand, studying the effects of service offshoring on white-collar employment in more than 100 U.S. occupations, Crinò (2010 b) concluded that (a) service offshoring increases employment in more skilled occupations relative to less skilled occupations; (b) at a given skill level, service offshoring penalizes offshorable occupations while benefiting less-offshorable occupations. However, evidence from European countries is mixed. Goos et al. (2010) found that offshoring was associated with reduced employment in offshorable occupations across 16 European countries as opposed to Crinò is (2010 a) finding that service offshoring has no effect on employment in Italian firms. Using occupational licensing as a shifter of offshoring costs, Criscuolo and Garicano (2010) found that an increase in service offshoring increased both wages and employment in less-offshorable service occupations (i.e., licensed occupations) in the UK.

Grossman and Rossi-Hansberg's theoretical framework includes wage implications and may partially relieve policymaker concerns over increased wage inequality due to offshoring in the U.S. labor market, but it does not address the core question of to what extent offshoring will affect labor demand (i.e., number of jobs). Goos et al. (2010) did find offshoring to be an explanatory factor affecting the conditional demand for labor in different occupations in their theoretical model and estimation, but other existing studies simply extend their empirical investigation to the effects of offshoring on wage or employment and provide some empirical evidences.

In this paper, Grossman and Rossi-Hansberg's (2008) offshoring model is generalized to include numerous tasks/skill levels (tasks correspond to specific occupations in this empirical framework) and investigate the effect of offshoring on occupational employment for ten major occupational groups (at 2-digit SOC level) in the U.S. labor market (see Table 3.2.1 and 3.2.2 for details of occupational groups). Using the CPSMORG (Current Popolution Survey Merged Outgoing Rotation Groups) data from year 1983 to 2011, analysis is conducted in two phases. First, the monotonic cubic spline interpolation method is used to estimate the offshoring cost functions for all ten occupational groups. The monotonic cubic spline interpolation method requires no specific functional form other than the assumption that offshoring costs are non-decreasing in the percentage of tasks being offshored. This nice property makes monotonic cubic spline interpolation method a perfect fit for this study because offshorability for one occupational group could largely differ from another. Next, a parametric method-nonlinear least squares (NLS)-is utilized for the five relatively offshorable occupational groups. Based on the monotonic cubic spline interpolation results, a cubic functional form is attached to the five relatively offshorable occupational groups to approximate their offshoring cost functions. Then, the number of jobs offshored and the offshoring percentage over the sample period for the five offshorable occupational groups are calculated.

Aside from a limited number of studies with primary information on offshoring activities (see for example, Crinò, 2010), researchers have used two alternative approaches to measure offshoring over time. The first approach is to approximate or infer offshoring activities using relevant information. For example, Ebenstein et al. (2013) use foreign affiliate employment of U.S. multinational firms as a measure capturing U.S. firms' offshoring activities. Criscuolo and Garicano (2010) use occupational licensing to infer the offshorability of an occupation in their study of offshoring of UK service sectors. Approximation of offshoring activities circumvents the issue of time-invariance of offshoring/offshorabilty index, but reliability of the approximation is unknown.

The second approach is to generate a time-invariant offshoring index based on firm offshoring activities. For example, Goos et al. (2010) construct an occupational offshorability index based on offshoring activities of 415 European firms between 2002 and 2008. Applying a time-invariant index assumes that the offshoring activities are either not influenced by the reduction of offshoring costs or that costs are constant over time. A time variant offshoring index is thus especially important when investigating the effect of offshoring over a relative long-time span. For example, the occupation of a radiologist would be considered as non-offshorable without the advancement in recent telecommunication technology which makes transformation of large image data a relatively costless task. An important contribution of this paper is to provide time-variant estimates of offshoring for more than 400 major U.S. occupations over the period of 1983 to 2011.

3.2 A Simple Structural Model of Offshoring

Inspired by empirical findings about the impact of characteristics of tasks on wage inequality and employment structure (e.g., Autor et al., 2003), Grossman and Rossi-Hansberg (2008) proposed a theoretical model of task offshoring to explain the impact of offshoring on the wage rates of different types of workers. In the Grossman and Rossi-Hansberg model, tasks are limited to only two types: low-skill and high- skill. Under a standard Heckscher-Ohlin set-up, Grossman and Rossi-Hansberg (2008) show how changing offshoring costs will affect the wage rates of low-skilled and high-skilled workers in the home country through static comparative analysis.

This research generalizes the analysis to include numerous tasks and link the concept of tasks to detailed occupations that are actually offshored. While the focal point of Grossman and Rossi-Hansberg (2008) is to decompose effects of offshoring on factor prices i.e., wage rates, this research focuses on exploring the effect of offshoring on employment levels in different occupations. To be consistent and comparable with Grossman and Rossi-Hansberg (2008), this research uses the term “task” instead of “occupation” in the structural model specification, but freely changes between these two in the remaining of this paper depending on the context. [2]

3.2.1 Model Specification

The production process requires many types of tasks and each type of task is denoted

[2] Each task corresponds to an occupation in our empirical framework.

by o. Producing one unit of a specific good involves a continuum of each type of task. Without loss of generality, the measure of each type of task can be normalized to one.

Firms in the home country produce many goods. The number of goods produced in the home country is assumed to be larger than the number of types of tasks.[3] All tasks are involved in order to produce one unit of specific good,[4] i.e., a_{oj} is the total amount of domestic factor *o* that would be needed to produce a unit of good *j* in the absence of any offshoring. Firms can undertake an o-type task either at home or abroad depending on the offshoring costs and the relative wage of task *o* between home and foreign country. An *o*-type task is indexed by $i \in [0, 1]$ and ordered in a manner such that the offshoring cost of task o, denoted by $t(i)$, is non-decreasing in t.

3.2.2 Model Derivation

As some tasks are more difficult to offshore than others, offshoring costs are assumed to be varying across different tasks and changing over time. Denote offshoring costs shifter as $\beta_{o,s}$ with subscript *o* indicating task type and *s* indicating time period. Let $w_{o,s}$ and $w^*_{o,s}$ be respectively the home and foreign wage of task *o*. Then the relative wage between home and foreign country of each task *o*, denoted by $\omega_{o,s}$, satisfies $\omega_{o,s} = \frac{w_{o,s}}{w^*_{o,s}}$ for all periods *s*.

---

[3] This assumption is to guarantee a unique solution to the factor price of each type of task given the price and production technology of each good.

[4] If the cost-minimizing demand for factor *o* is zero, the *o*-type task will be missing in the production process.

Following Grossman and Rossi-Hansberg (2008)'s formulation, $I_{o,s}$, the equilibrium marginal task *o* performed at home (or the cutoff point of task *o* at equilibrium) in period *s* in each industry is determined by the following condition such that wage savings exactly balance the offshoring cost of task *o*:

$$w_{o,s} = w^{*}_{o,s}\beta_{o,s}t\left(I_{o,s}\right) . \tag{3.1}$$

Then by my relative wage assumption $\omega_{o,s} = \frac{w_{o,s}}{w^{*}_{o,s}}$,

$$t\left(I_{o,s}\right) = \frac{\omega_{o,s}}{\beta_{o,s}} = \rho_{o,s} , \tag{3.2}$$

where $\rho_{o,s}$ denotes the equilibrium offshoring costs, which depends on the ratio of relative wage $\omega_{o,s}$ and the offshoring cost shifter $\beta_{o,s}$ at each period s. Given that $t(\cdot)$ is an increasing function in $I_{o,s}$, a higher proportion of task *o* will be moved offshore as $I_{o,s}$ increases. As $I_{o,s}$ is the cutoff point of the marginal task *o* performed at home country, $\rho_{o,s}$ precisely captures the offshoring decisions made by home firms.

Denote L_o the initial total employment of occupation *o* at home country without offshoring and $L_{o,s}$ the employment of occupation *o* in period s with offshoring, which is observed in data, then $L_{o,s}$, can be calculated as following:

$$L_{o,s} = \left(1 - \ I_{o,s}\right) \cdot L_o, \tag{3.3}$$

where $1 - I_{o,s}$ indicates the fraction of o-type tasks that are performed at home.

Under the perfect competitive assumption, the price of any good j is equal to the unit cost of production (if a positive quantity of the good is produced):

$$p_j = \sum_o w_{o,s}\Omega(\mathrm{I}_{o,s})a_{oj}(\cdot), \qquad (j > o) \tag{3.4}$$ [5]

where, the arguments in the function for the factor intensity a_{oj} are the relative costs of the various sets of tasks when they are located optimally with offshoring,

and $$\Omega(\mathrm{I}_{o,s}) = 1 - \mathrm{I}_{o,s} + \frac{\int_0^{I_{o,s}} t(i)di}{t(I_{o,s})}. \tag{3.5}$$

In other words, $\Omega(\mathrm{I}_{o,s})$ consists of two parts, $1 - \mathrm{I}_{o,s}$ (the proportion of tasks that remains in home country) and $\frac{\int_0^{I_{o,s}} t(i)di}{t(I_{o,s})}$ (the proportion of tasks conducted in foreign country expressed in equivalent home-country factor employment).

As $I_{o,s} = \frac{L_O - L_{O,S}}{L_O} = 1 - \frac{L_{O,S}}{L_O}$ is a function of L_O, $\Omega(\mathrm{I}_{o,s})$ is a function of L_O.

Since the number of the goods is larger than the number of factors $(j > o)$, factor prices$\left(w_{o,s}\Omega(\mathrm{I}_{o,s})\right)$ can be uniquely determined and solved from the systems of equations (3.4). That is,

$$w_{o,s}\Omega(\mathrm{I}_{o,s}) = c_o, \tag{3.6}$$

---

[5] Equivalent to Equation (3) in Section I, Grossman and Rossi-Hansberg (2008). For detailed derivation, please refer to Grossman and Rossi-Hansberg (2008).

where c_o depends on the prices p_j and all production technologies of all goods produced in home country. Identity (3.6) is the key equation in identifying the equilibrium cutoff point of offshoring percentage $(\mathrm{I}_{o,s})$ of task o, offshoring cost $t(i)$ as well as constant c_o. Section 3.2.3 explains estimation of Equation (3.6).

3.2.3 Model Interpretation

Although the Grossman and Rossi-Hansberg (2008) model is static, it can be interpreted with some dynamics within each period. Given the wage differential between the home and foreign country, the equilibrium cutoff point of offshoring $I_{o,s}$ is determined by Equation (3.1) at the beginning of period s, which automatically determines the domestic labor demand for task o (in Equation (3.3)). By the zero-profit condition under perfect competition, the new wage $w_{o,s}$ for task o at the end of period s in the home country is obtained by solving Equation (3.4) (or equivalently Equation (3.6)). If the new wage $w_{o,s}$ is higher (or lower) than the starting wage in period s, the firm in the home country increases (or decreases) offshoring until it reaches its new equilibrium cutoff point at the end of period s that we observe in the data. The same process repeats in all periods.

By this interpretation, it is explicitly assumed the wage and employment observed in our data set are equilibrium wage and employment at the end of each period, which are both driven by offshoring. Then by estimating Equation (3.6), we can identify the offshroing cost function $t(i)$[6] and the initial employment without offshoring for each task o.

[6] However, the offshoring cost function $t(i)$ can only be identified up to a constant scale because multiplying a scalar to $t(i)$, Equation (3.6) still holds.

3.3 Estimation Framework and Method

3.3.1 The Empirical Framework

To estimate Equation (3.6), take logarithm and reorder, which leads to,

$$lnw_{o,s} = -ln\Omega(\mathrm{I}_{o,s}) + lnc_o = lnc_o - ln\Omega(\mathrm{I}_{o,s}). \qquad (3.7)$$

As $\Omega(\ \mathrm{I}_{o,s})$ is a function of the observed variable $L_{o,s}$, unobserved parameters L_o and the offshoring cost function $t(\cdot)$, standard linear estimation methods which can only estimate unknown parameters but not unknown functions are not applicable.

Further denote $y_{o,s} = \ lnw_{o,s}$, $x_{o,s} = L_{o,s}$. Then the conditional mean of $y_{o,s}$ can be specified as

$$E(y_{o,s}|x_{o,s}) = m(x_{o,s}, \boldsymbol{\theta_0}) = lnc_o - ln\Omega\left(\ \mathrm{I}_{o,s}(L_{o,s}, L_{o,}t(\cdot))\right) \qquad (3.8)$$

Where $\boldsymbol{\theta_0} = (L_o, c_o\ , t(\cdot))$ consists of two parameters and one function to be identified. Since $\boldsymbol{\theta_0}$ contains the offshoring cost function that cannot be directly estimated, I need to parameterize $t(\cdot)$ in order to proceed to estimate $t(\cdot)$ together with the other two parameters.

No specific structure except the monotonicity of $t(i)$ (i.e., $t(i)$ is non-decreasing in i) is assumed in the Grossman and Rossi-Hansberg (2008) framework. Hence, using a parametric estimation method and attaching any specific functional form to the offshoring cost function $t(i)$ for all ten occupational groups in empirical estimation will likely result in misspecification problems. Instead the non-parametric cubic spline method, in particular, the monotonic cubic spline interpolation method is adopted to approximate the offshoring cost function $t(i)$.

Once parameterization of $t(\cdot)$ is resolved, estimation of equation (3.8) becomes a standard non-linear estimation problem. The NLS estimators

$$\boldsymbol{\theta} = \min_{\theta \in \Theta} N^{-1} S^{-1} \sum_{o=1}^{N} \sum_{s=1}^{S} \left\{ y_{o,s} - m\left(\boldsymbol{x}_{o,s}, \boldsymbol{\theta}\right) \right\}^2 \quad (3.9)$$

minimize the sum of least squared residuals of the sample average and should solve the sample minimization problem if the true parameters $\boldsymbol{\theta_0} = argmin_{\theta \in \Theta} E\{[y - m(x, \boldsymbol{\theta})]^2\}$ solve the population minimization problem.

Ideally we would estimate Equation (3.8) occupation by occupation to identify the initial employment without offshoring L_o at home country, the constant parameter c_o and the set of parameters for each occupation *o* in the parameterized offshoring cost function $t(i)$. Due to data restrictions, [7] the individual occupations are grouped into ten broad occupational groups for pre- and post-2000 periods respectively and these groups are used as the basis to estimate Equation (3.8). [8]

3.3.2 Application of Monotonic Cubic Spline Interpolation Method

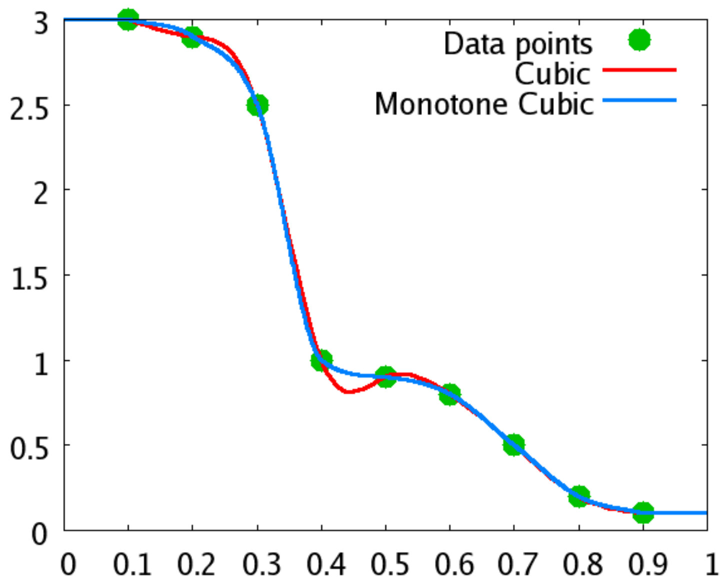
A two-step monotonic cubic spline interpolation procedure is used to estimate $\widehat{\boldsymbol{\theta}} = \left(L_{\boldsymbol{o}}, c_{\boldsymbol{o}}, t(i)\right)$ for the ten occupational groups based on the algorithm of monotonic cubic

[7] See data description for details.

[8] To distinguish, subscript $\boldsymbol{o}$ (bold *italic*) is used to represents an occupational group in the remaining of this paper.

spline interpolation developed by Wolberg and Alfy (1999, 2002). Figure 3.1 illustrates an example of monotonic cubic spline: the interpolating cubic spline passing through its control points is smooth and monotonic. While it is not yet often used in the field of economics, monotonic cubic spline interpolation is a well-developed method and widely used in numerical and statistical data analysis to solve engineering problems. The most compelling reason for the use of cubic polynomials is the property of twice differentiable continuity, which guarantees continuous first and second derivatives across all intervals. The goal of cubic spline interpolation is to determine the smoothest possible curve that passes through designated control points while simultaneously preserving the property of piecewise monotonicity within each interval.

Figure 3.1: An Example of Cubic Spline Interpolation



Source: Wikipedia (http://en.wikipedia.org/wiki/Monotone_cubic_interpolation).

(For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.)

The algorithm of Wolberg and Alfy (2002) is adopted in the first step. The algorithm itself consists of two parts designated WAA Step-1 (abbreviation of Wolberg and Alfy Algorithm) and WAA Step-2 to distinguish from the overall two-step interpolation procedure and avoid confusion. The WAA Step-1 attempts to find a twice continuously differentiable cubic spline which minimizes the modified second derivative discontinuity in the spline.[9] If a twice continuously differentiable cubic spline exists, the WAA Step-2 is then employed to obtain the optimal twice continuously differentiable cubic spline by

[9] Definition of second derivative discontinuity: $\sum_i [f''(x_i^-) - f''(x_i^+)]^2$. Definition of modified second derivative discontinuity: summation of second derivative difference is non-negative, i.e., $\sum_i [f''(x_i^-) - f''(x_i^+) + K] \geq 0$, where K satisfies $f''(x_i^-) - f''(x_i^+) + K \geq 0$ for any arbitrary i. The reason to use modified second derivative discontinuity is to turn the objective function into a linear function so that linear programming can be applied. See Wolberg and Alfy (2002) for details.

computing the integral of the spline curvature. If not, the best first differentiable cubic spline is obtained in the WAA Step-1 and the WAA Step-2 is canceled.

To estimate $\boldsymbol{\theta}$, the offshoring percentage interval $i \in [0,\ 1]$ is partitioned into ten equal sub-intervals, representing the percentage increment of i being offshored. The WAA Step-1 and WAA Step-2 are applied to locally approximate the offshoring cost function $t(i)$ and obtain the monotonic cubic spline interpolation for each occupational group. The interpolated offshoring cost function is then used to calculate $\Omega\left(\ \mathrm{I}_{o,s}\left(L_{o,s}, L_{o,} t(\cdot)\right)\right)$ in Equation (3.9). Then the optimal estimators of $\widehat{\boldsymbol{\theta}}$ is obtained by minimizing the non-linear least square errors by iterations. $\widehat{\boldsymbol{\theta}}$ is a vector containing 13 estimators. They are estimator of the initial employment of occupation o at home country without offshoring $\widehat{L}_{\boldsymbol{o}}$, estimator of the constant parameter $\hat{c}_{\boldsymbol{o}}$ and the set of estimators for parameterized offshoring cost function $t(i)$, which corresponds to 11 control points that portioned $i \in [0,\ 1]$ into ten equal sub-intervals.

3.3.3 Estimating Offshoring Cost Functions for the Ten Major Occupational Groups

Implementation of the monotonic cubic spline approximation to estimate offshoring cost functions for the ten major occupational groups requires updating the initial values of the cost function $t(i)$ at each control point of i. [10] Hence initial starting values for $t(\cdot)$ must be obtained. Blinder and Krueger (2009) provide values for offshorability in major occupational groups[11] as the starting point to differentiate relatively offshorable occupations

[10] The 11 control points of i are 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8., 0.9, 1.

[11] See Table 2, Column 5, titled Externally-Coded Percent Offshorable in Blinder and

from relatively non-offshorable occupations.[12] Based on their externally-coded estimates, the ten occupational groups are divided into two broad categories: Offshorable Groups and Non-offshorable Groups (Table 3.1).

Table 3.1: Offshorablility in Major Occupational Groups

Rank of Offshorablity	Occupational Group (Externally-coded Offshorable Percentage)
Offshorable Groups	
1	G9: Production occupations[13] (80.7%)
2	G5: Office and administrative support occupations (41.2%)
3	G2: Professional and related occupations (20.5%)
4	G4: Sales and related occupations (17.8%)
5	G1: Management, business, and financial operations occupations (16.4%)
Non-offshorable Groups	
6	G6: Farming, fishing, and forestry occupations (0.0%)
6	G7: Construction and extraction occupations (0.0%)
6	G10: Transportation and material moving occupations (0.0%)
7	G3: Service occupations (0.7%)
8	G8: Installation, maintenance, and repair occupations (1.3%)

Notes: Prepared by authors based on the externally-coded offshorable percentage (Column 2, Table 2) in Blinder and Krueger (2009).

Adjustment of employment size for each occupation within an occupational group is necessary before the monotonic cubic spline approximation is applied to estimate the offshoring cost functions for the ten major occupational groups. There are large between-occupation variations in employment within a same occupational group (Table 3.4).

Krueger (2009).

[12] There are sharp disagreements between self-classified and externally coded offshorability for some occupational groups. This research uses the externally-coded offshorability by professionals as the criterion to divide offshorable and non-offshorable groups.

[13] For the purpose of simplicity and comparability with Blinder and Krueger (2009)'s results, only post-2000 occupation titles are used to indicate occupational groups in the main text unless otherwise specified.

However, by estimating Equation (3.8) at the basis of occupational groups, the to-be-identified parameter $L_{\boldsymbol{o}}$ (i.e., the initial total employment without offshoring) is implicitly assumed to be same for all occupations within a group. This is a relatively strong assumption for the ten occupational groups with large between-occupation variations in employment within each occupational group. In order to identify $L_{\boldsymbol{o}}$ and obtain a meaningful $\hat{L}_{\boldsymbol{o}}$ for each occupational group, this study adjusts employment size to make employment size for each occupation relatively homogenous within an occupational group.[14]

As $\hat{L}_{\boldsymbol{o}}$ (i.e., estimated total employment without offshoring) is heavily dependent on the within-group variations of adjusted occupational employment $\tilde{L}_{o,s}$ between different occupations, offshoring percentage is restricted to not exceed 10% of the maximum $\tilde{L}_{o,s}$ (i.e., the maximum adjusted employment of all occupations across all years in the sample period) in each occupational group. In other words, the estimated $\hat{L}_{\boldsymbol{o}}$ is restricted to $\hat{L}_{\boldsymbol{o}} \leq 1.1 * max(\tilde{L}_{o,s})$. This restriction is also used as one stopping criterion for iterations when applying monotonic cubic spline interpolation to approximate the offshoring cost functions for the ten occupational groups.

[14] Detailed adjustment method of employment size is discussed in Section 3.4 after introducing the data set.

3.3.4 Estimating Number of Jobs Offshored and Offshoring Percentage for the Five Relatively Offshorable Occupational Groups

After estimation of the offshoring cost functions for the ten major occupational groups, this analysis is focused on the five relatively offshorable occupational groups in Table 3.1. Nonlinear least squares (NLS) with a specific cubic functional form $t(i) = ai^3 + bi^2 + di + e$ is employed to re-estimate the offshoring cost functions for the five relatively offshorable occupational group. The number of jobs offshored and the offshoring percentage by detailed occupation in pre- and post-2000 sample period are calculated after estimation of the cubic offshoring cost function. There are a few reasons to focus on the relatively offshorable occupational groups. First, offshoring cost is relatively tractable because fluctuations of employment at offshorable occupations over time observed in data reflect the change of offshoring costs. Second, factors (e.g., technology, institutional restructuring) that could potentially affect the occupational employment except offshoring are not controlled in this study. In other words, changes of employment over time are assumed to be purely attributable to offshoring in this framework. While this is a strong assumption, it is more realistic for the relatively offshorable occupations which are primary focus of this study.

To calculate the number of jobs offshored and the offshoring percentage for the five offshorable occupational groups over the pre- and post-2000 sample period, this study uses $\hat{L}_{\boldsymbol{o}}$, which is estimated from the adjusted employment size $\tilde{L}_{o,S}$ of each occupation from the two-step cubic spline interpolation method, to recover $\tilde{L}_{\boldsymbol{o}}$, the unadjusted initial employment without offshoring for each occupational group by reversing the adjustment method.

Different scenarios are applied when re-estimating the cubic offshoring cost functions using NLS, calculating the number of jobs offshored and offshoring percentage for the five offshorable occupational groups. Based on Blinder and Krueger's (2009) estimates for offshorable occupational groups (re-organized in Table 3.1), the 20% scenario is chosen as a benchmark case for all five groups because the externally coded offshorability are relatively close to 20 percent (Group 1, Production occupations, 16.4%; Group 2, Professional and related occupations, 20.5%; Group 4, Sales and related occupations, 17.8%). In the 20% scenario, the offshoring percentage is assumed to not exceed 20 percent of the maximum $\tilde{L}_{o,s}$, i.e., the estimated $\hat{L}_{\boldsymbol{o}} \leq 1.2 * max\left(\tilde{L}_{o,s}\right)$. The maximum offshoring percentage is then gradually relaxed to 40 percent (externally coded offshorable percentage is 41.2 percent for Group 5, Office and Administrative Support Occupations) and 80 percent (externally coded offshorable percentage is 80.7% for Group 9, Production Occupations) for all five offshorable groups.

3.4 Data Description and Adjustment

The CPSMORG (Current Population Survey Merged Outgoing Rotation Groups) data from years 1983 to 2011 are used to implement the two-step monotonic cubic spline interpolation procedure. The data are discontinuous due to a complete switch in the occupational and industrial classification system in CPS (Current Population Survey) in 2003.[15] This substantial change in the composition of detailed occupations between the 1980 and 2002 occupation codes makes linking data by occupation codes impossible. Hence, the sample is divided into two periods: pre-2000 (1983-1999) and post-2000 period (2000-2011) to conduct analysis at occupational level.

[15] Years 2000-2002 are dual-coded in both 1980 and 2002 census classifications systems.

Observations for individuals with age less than 18 and or more than 65 are dropped from the sample to maintain focus on the labor force. Hourly wage series for each individual is created following Schmitt 2003 and inflated by 2000 CPI index to obtain the real hourly wage. Wage and employment are aggregated to occupation level based on 1980 census codes for the pre-2000 period and based on 2002 census codes for the post-2000 period. CPS earning weights are used to obtain occupational hourly wage while CPS final weights are used to obtain occupational employment during aggregation. To maintain balanced panels for both the pre- and post-2000 periods, occupations not present in all years of each analysis period were omitted from the data set. After aggregation, there are 486 occupations in the pre-2000 period and 460 occupations in the post-2000 period (Table 3.2.1 and 3.2.2).

As mentioned earlier, by estimating offshoring cost functions by occupational groups L_o is implicitly assumed to be same for all occupations within an occupational group. But the large between-occupation variations in employment within an occupational group is not in favor of this assumption. Several adjustments are made to reduce between-occupation variations and homogenize the employment size within each occupational group.

For both pre- and post-2000 sample period, mean employment for each occupation is calculated and a median employment for all occupations within an occupational group is obtained. Relative employment size for each occupation is mean employment of each occupation by this occupational group median employment.[16] Finally, the adjusted employment for each occupation in each year $\tilde{L}_{o,s}$ is observed employment $L_{o,s}$ divided by the relative employment size of each occupation. The adjusted employment for each

---

[16] If there are even-numbered groups within an occupational group, we use the larger of the two medians as the denominator.

occupation $\tilde{L}_{o,s}$ is used in the monotonic cubic spline interpolation to approximate the offshoring cost function.

The estimated $\hat{L}_{\boldsymbol{o}}$ largely depends on the maximum or minimum value of the adjusted employment $\tilde{L}_{o,s}$ within each occupational group. The estimated $\hat{L}_{\boldsymbol{o}}$ is likely to be misleadingly inflated if there are extreme values of $\tilde{L}_{o,s}$ within an occupational group. Hence, occupations with observations falling in the upper and lower five percentile of the adjusted employment are dropped to further homogenize the employment size of occupations within each occupational group. Table 3.4 summarizes the employment size variations for each occupational group before and after adjustment for pre- and post-2000 periods respectively. After adjustment, the mean and median employment size within each occupational group are quite close. The between-occupation employment variations within an occupational group are largely reduced.

3.5 Results and Discussion

3.5.1 Offshoring Costs for the Ten Major Occupational Groups

The 11 point estimates of the parameterized offshoring cost functions from the monotonic cubic spline interpolation method for the ten major occupational groups are summarized in Table 3.5. A corresponding interpolated offshoring cost function $t(i)$ for each occupational group are plotted in Figure 3.2.1 and Figure 3.2.2 for the pre- and post-2000 periods respectively. The estimated $\hat{c}_o$ and $\hat{L}_o$ are reported in Table 3.6.

One issue to be emphasized in front is that any direct comparison between pre- and post-2000 periods is not feasible although results for the pre- and post-2000 periods are sometimes displayed in parallel. As mentioned earlier, compositions of occupations within each occupational group for pre- and post-2000 periods are completely different. Consequently, the estimated $\widehat{L}_O$ in pre-2000 period is not comparable with the estimated $\widehat{L}_O$ in post-2000 period due to this occupational composition difference. Nonetheless, results from these two sample periods are consistent and some general patterns can be observed. Estimated offshoring cost functions indicate an effect of economies of scale in offshoring. The offshoring cost increases in the first ten percent of offshoring and then decreases as more jobs offshored.[17] Among the ten occupational groups, Group 1 (Management, business, and financial operations occupations), Group 2 (Professional and related occupations), Group 4 (Sales and related occupations), Group 5 (Office and administrative support occupations) and Group 9 (Production occupations) have relatively lower costs at any given level of offshoring percentage i in both the pre- and post-2000 periods. In particular, production occupations in Group 9, which are commonly regarded to contain the most impersonal and/or routine tasks and easiest to offshore, have the lowest offshoring costs when the offshoring percentage is below 40 percent (Table 3.4). The estimated offshoring cost for production occupations has a sharp increase when offshoring moves from the first 40 percent to 50 percent in the pre-2000 period, and from the first 30 percent to 40 percent in the post-2000 period. The remaining five occupational groups, Group 3 (Service occupations), Group 6 (Farming, fishing, and forestry occupations),[18]

[17] The partition of 10 subintervals is arbitrary. But increasing the numbers of subintervals does not alter the result because the nice monotonic property of monotonic cubic spline interpolation within each interval.

[18] Group 6 has low offshoring cost (small point estimates) in the first 30 percent of offshoring due to few observations between interval 0.0 and 0.3.

Group 7 (Construction and extraction occupations), Group 8 (Installation, maintenance, and repair occupations) and Group 10 (Transportation and material moving occupations), have relatively higher offshoring costs.

The rank of offshorability in this study based on estimated offshoring costs for both pre- and post-2000 periods is different from the externally coded offshorability of Blinder and Krueger (2009) based on individual telephone survey in 2008, but most results are consistent with them. Blinder and Krueger (2009) found Group 6 (Farming, fishing, and forestry occupations), Group 7 (Construction and extraction occupations) and Group 10 (Transportation and material moving occupations) to be the least offshorable. This study identified farming, fishing, and forestry occupations (Group 6), construction and extraction occupations (Group 7), and service occupations (Group 3) with the highest offshoring costs while transportation and material moving occupations in Group 10 with the second highest offshoring cost.

3.5.2 NLS Results for the Five Relatively Offshorable Occupational Groups

The estimated coefficients of the cubic offshoring cost function together with $\hat{c}_O$ and $\hat{L}_O$ by NLS method under three different scenarios are reported in Table 3.5. Corresponding offshoring cost functions $t(i)$ of the five relatively offshorable groups are displayed respectively in Figure 3.3.1 through Figure 3.3.5. Unlike the monotonic cubic spline interpolation method, it is difficult to directly compare the estimated cubic offshoring cost functions among different occupational groups within the same scenario, or the same occupational group among three different scenarios given the fact that the single cubic functional form attached to all five relatively occupational groups cannot be uniquely identified because there is only one moment condition (i.e., Eq. 3.6) available in the structural model.

The number of jobs offshored and the offshoring percentage are calculated based on the estimated $\widehat{L}_o$ for the five relatively offshorable groups in pre- and post-2000 periods are summarized in Table 3.8.1 through Table 3.8.5. First, the initial total employment for each occupation o in each year s within an offshorable occupational group is recovered by multiplying the relative employment size of each occupation to its corresponding $\widehat{L}_{\boldsymbol{o}}$ estimated for each occupational group. The number of jobs offshored at each occupation o in each year s is then the difference between the recovered initial total employments of occupation o and $L_{o,s}$ observed in data. The offshoring percentage is then obtained using the number of jobs offshored divided by the initial total employment without offshoring.

Both the number of jobs offshored and offshoring percentage increase as the maximum offshoring capacity increases from 20% scenario to 80% scenario. Production occupations in Group 9 have been consistently increasing over time in both pre- and post-2000 periods under all three scenarios. Under the 20% scenario that maximum 20 percent of production occupations are offshorable, the offshoring percentage for production occupations increases from 36.5 to 46.3 percent in the pre-2000 period and increases from 36.1 to 48.5 percent in the post-2000 period. Under the 40% scenario, offshoring percentage for production occupations increases from 45.6 to 54.0 percent in the pre-2000 period and increases from 45.2 to 55.9 percent in the post-2000 period. Under the 80% scenario, offshoring percentage for production occupations increases from 57.4 to 64.0 percent in the pre-2000 period and from 52.9 to 62.1 percent in the post-2000 period, which are less than the estimated 80.7 percent by Blinder and Krueger (2009).

Changes in the offshoring percentage for the five relatively offshorable occupational groups over the two sample periods are depicted in Figure 3.4.1 through Figure 3.4.5

additionally. Offshoring percentage for sales and related occupations in Group 4 and office and administrative support occupations in Group 5 have been gradually increasing over time in the post-2000 period. On the other hand, for management, business and financial operations occupations (Group 1) and professional and related occupations (Group 2), offshoring percentage actually has decreased over time.

In addition, using externally coded offshorability estimated for the five offshorable groups from Blinder and Krueger (2009) (reorganized in Table 3) as a criterion, results of the 20% scenario for Group 1, Group 2 and Group 4, the 40% scenario for Group 5, and the 80% scenario for Group 9 to are selected to make comparisons among groups. This comparison shows that occupations in sales and related occupations in Group 2 are the least offshorable among the five offshorable occupational groups followed by the management, business and financial operations occupations and professional and related occupations.

3.6 Conclusion

This research generalizes the Grossman and Rossi-Hansberg (2008) offshoring model to include numerous tasks/skill levels. This generalization allows a possible and direct linkage between the theoretical task offshoring model and occupational data that can be aggregated from the CPSMORG (Current Population Survey Merged Outgoing Rotation Groups) data from year 1983 to 2011. Empirical investigation of the effect of offshoring on occupational employment for ten major occupational groups (at 2-digit SOC level) in the U.S. labor market is conducted by estimating their offshoring cost functions using a non-parametric monotonic cubic spline interpolation method. Based on the estimated offshoring costs, five relatively offshorable occupational groups are identified including production occupations, office and administrative support occupations, sales and related occupations, professional and related occupations, and management, business, and financial operations occupations.

Motivated by the practical issue of difficulty in obtaining a time-variant offshoring/offshorability index faced by majority empirical studies interested in identifying the effect of offshoring, this study calculates the number of jobs offshored and the offshoring percentage under the NLS method for the five relatively offshorable occupational groups under different scenarios. Calculated offshoring percentage provides time-variant offshoring indices for more than 300 major detailed occupations in these five relatively offshorable groups that can be employed in other empirical studies.

The results of this research indicate that offshoring percentage for each occupational group may vary under different scenarios, but the evolution pattern is consistent. Production occupations are the most offshorable while sales and related occupations are the least offshorable among all five offshorable occupational groups under all three scenarios. The offshoring percentage for production occupations has been increasing in both pre- and post-2000 periods while the offshoring percentages for professional and related occupations, and management, business, and financial operations occupations have been decreasing over time.

APPENDIX

Table 3.2.1: Major Occupational Groups in Pre-2000 Period (1983-1999)

Group	1980 Census Codes	Occupation Title	Number of Occupations
1	003-037	Managerial and professional Specialty occupations	24
2	043-199 203-235	Professional specialty occupations Technical occupations	126
3	403-469	Service occupations	42
4	243-285	Sales occupations	23
5	303-389	Administrative support occupations	55
6	473-499	Farming, forestry, and fishing occupations	19
7	553-599 613-617	Construction trades Extractive occupations	35
8	503-549	Mechanics and Repairers	27
9	633-699 703-799	Precision Production Occupations Operators, fabricators, and laborers	99
10	803-889	Transportation and Material Moving Occupations	39
Total			486

*Notes: Occupational group information is obtained from (http://usa.ipums.org/usa/volii/98occup.shtml), but reorganized and reordered by author to be comparable with occupational groups in post-2000 period.

Table 3.2.2: Major Occupational Group in Post-2000 Period (2000-2011)

Group	2002 Census Codes	Occupation Title	Number of Occupations
1	0010-0950	Management, business, and financial operations occupations	42
2	1000-3540	Professional and related occupations	107
3	3600-4650	Service occupations	57
4	4700-4960	Sales and related occupations	17
5	5000-5930	Office and administrative support occupations	50
6	6000-6130	Farming, fishing, and forestry occupations	8
7	6200-6940	Construction and extraction occupations	36
8	7000-7620	Installation, maintenance, and repair occupations	34
9	7700-8960	Production occupations	75
10	9000-9750	Transportation and material moving occupations	34
Total			460

*Notes: Occupational groups are equivalent to those grouped at 2-digit SOC level.

Table 3.3: Occupational Employment Size[1] Variation

Pre-2000 Period: 1983-1999								
Group	Before Adjustment $\left(\overline{L_{o,s}}\right)$				After Adjustment[2] $\left(\overline{\tilde{L}_{o,s}}\right)$			
	Max	Min	Median	Mean	Max	Min	Median	Mean
1	5,179,799	10,103	234,885	509,254	336,438	168,371	257,054	255,055
2	1,829,530	2,439	60,309	155,601	85,892	35,756	60,110	60,309
3	2,463,471	11,309	218,522	433,122	295,584	147,323	224,454	222,077
4	3,470,040	16,358	239,788	624,399	296,549	197,353	247,542	245,897
5	4,064,116	4,676	186,986	368,996	260,335	117,224	188,326	186,986
6	1,173,238	1,894	39,952	154,177	61,434	22,577	40,062	39,952
7	1,103,129	3,192	42,083	135,112	55,790	27,795	42,113	42,083
8	746,818	3,043	101,811	161,553	131,050	76,843	101,989	101,811
9	1,409,946	2,775	42,481	121,924	59,714	25,258	42,741	42,689
10	2,780,569	3,098	87,280	289,827	54,206	120,427	87,752	7,280
Post-2000 Period: 2000-2011								
Group	Before Adjustment $\left(\overline{L_{o,s}}\right)$				After Adjustment $\left(\overline{\tilde{L}_{o,s}}\right)$			
	Max	Min	Median	Mean	Max	Min	Median	Mean
1	2,402,506	8,412	180,133	359,564	231,030	138,907	181,858	181,400
2	2,819,706	2,485	80,408	202,254	106,722	55,535	81,007	81,271
3	2,274,862	4,214	115,561	377,052	145,055	81,288	116,166	115,561
4	3,548,378	34,789	325,546	851,402	386,948	250,422	323,554	325,546
5	3,507,671	5,690	150,514	388,964	214,546	112,681	158,846	159,529
6	868,469	2,253	20,680	154,201	29,347	11,050	20,308	20,680
7	1,440,582	3,107	45,802	182,013	66,277	27,194	46,395	46,183
8	766,161	2,940	48,820	126,207	65,622	31,711	49,031	49,308
9	1,185,664	3,646	40,289	113,951	60,173	21,929	39,664	40,289
10	3,089,586	3,943	52,905	257,220	76,506	34,108	52,943	53,229

Notes: 1. For each occupation, employment is averaged across years within sample period.
2. Occupations with employment falling in the upper and lower five percentile are dropped after adjustment.

Table 3.4: Point Estimates of Parameterized Offshoring Cost Function $t(i)$ from Cubic Spline Interpolation Method for Ten Major Occupational Groups

Group	Value of $t(i)$ at Control Points i Pre-2000: 1983-1999										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	5.33	3.38	11.5	25.8	38.4	49.6	70.3	77.7	95.6	137.3	202.1
2	1.12	2.65	10.3	15.4	32.5	71.9	111.5	137.7	130.7	124.4	108.8
3	14.8	6.05	6.04	22.2	49.3	173.7	252.7	372.5	297.8	452.8	259.4
4	0.01	6.60	12.6	20.8	40.5	56.1	87.8	120.5	185.8	167.5	144.5
5	10.2	1.18	5.36	24.5	41.2	45.8	55.2	83.1	107.3	136.5	145.7
6	7.44	12.0	19.6	40.0	51.1	101.8	119.0	170.5	249.8	369.4	505.1
7	4.19	9.27	17.4	25.2	75.4	52.3	92.7	136.0	193.1	307.4	448.7
8	0.40	0.20	16.4	16.3	73.5	110.6	128.8	162.2	265.8	356.5	551.8
9	8.24	0.14	1.07	7.34	18.9	53.8	103.5	132.0	151.5	182.6	166.0
10	11.7	9.35	19.1	42.8	74.3	182.2	279.1	424.0	551.8	766.9	1006.3
	Value of $t(i)$ at Control Points i Post-2000 Period: 2000-2011										
1	2.71	6.35	12.5	20.4	30.5	42.6	58.43	77.38	91.6	121.1	188.1
2	5.98	2.02	4.88	9.74	31.9	113.0	147.2	197.9	227.4	175.7	460.9
3	20.3	18.3	57.8	73.5	142.5	233.0	180.9	404.9	543.5	747.3	1076.9
4	0.00	0.38	13.0	29.3	29.6	53.2	61.3	86.4	104.5	120.1	184.9
5	3.18	1.04	0.88	38.3	35.5	47.0	64.4	85.4	101.7	133.4	168.8
6	20.4	23.5	25.4	35.8	53.7	77.8	109.2	168.0	247.8	222.3	262.3
7	2.21	27.8	26.6	40.8	23.2	82.2	116.5	203.8	264.8	363.0	543.6
8	5.67	7.27	2.13	18.2	34.6	40.9	52.5	92.0	144.7	182.2	269.8
9	0.54	0.09	0.90	41.1	20.7	69.8	103.4	95.5	85.6	117.3	134.2
10	17.9	11.1	9.83	53.7	80.0	99.9	132.7	170.6	121.8	195.2	241.1

Notes: 1. No other control variables are included in the model.
2. The upper bound is set that offshoring cannot exceed the 10% of the maximum employment of all occupations across all years within each group.
3. The maximum iterations is 500 times.
4. Initial value is adopted from the first-round cubic spline interpolation results without dropping any observations. See Table 3.A for details of initial value.

Table 3.5: Estimates of $\hat{L}_O$, $\hat{c}_O$ by Occupational Groups from Cubic Spline Interpolation Method for Major Ten Occupational Groups

Group	Pre-2000 Period (1983-1999)		Post Post-2000 Period (2000-2011)	
	$\hat{L}_O$	$\hat{c}_O$	$\hat{L}_O$	$\hat{c}_O$
1	370,034	5.86	245,310	7.00
2	89,912	6.69	110,980	6.51
3	323,423	2.72	152,963	3.82
4	314,449	3.74	425,643	5.05
5	284,791	3.27	225,284	3.77
6	67,578	1.85	29,860	3.43
7	61,369	4.21	69,592	4.98
8	137,680	5.05	69,206	5.50
9	62,628	3.40	66,190	3.13
10	367,550	5.95	84,049	4.20

Notes: 1. No other control variables are included in the model.
2. The upper bound is set that offshoring cannot exceed the 10% of the maximum employment of all occupations across all years within each group.
3. The maximum iterations is 500 times.
4. Initial value is adopted from the first-round cubic spline interpolation results without dropping any observations. See Appendix Table 3.A for details of initial value.

Table 3.6: NLS Estimates of Cubic Offshoring Cost Function $\hat{L}_o$,$\hat{c}_o$ for the Five Relatively Offshorable Occupational Groups

Group	Pre-2000 (1983-1999)						Post-2000 (2000-2011)					
	Coefficients of Cubic $t(i)$				$\hat{L}_o$	$\hat{c}_o$	Coefficients of Cubic $t(i)$				$\hat{L}_o$	$\hat{c}_o$
	$\hat{a}$	$\hat{b}$	$\hat{d}$	$\hat{e}$			$\hat{a}$	$\hat{b}$	$\hat{d}$	$\hat{e}$		
	20% Scenario											
1	-3875	-4341	-1678	348.2	403,726	5.03	16397	19786	5933	-1380	277,236	5.63
2	-4062	-4267	-2823	549	103,070	4.94	-8.77	-9.30	-3.55	0.81	128,066	5.24
4	-2085	-2335	-896	195	355,858	2.88	-1733	-1916	-754	160	464,337	4.30
5	-16.91	-17.49	-7.04	1.54	312,402	2.85	2272	2582	972	-205	257,455	3.09
9	-5686	-6249	-2337	624	71,622	2.72	17.92	20.86	7.23	-1.63	72,208	3.00
	40% Scenario											
1	5136	5783	2149	-922.8	471,014	4.02	1578	5273	218	-407	323,442	4.46
2	2021	1801	966	-394	120,249	3.72	-1549	-1910	-723	338	149,410	3.99
4	-0.001	-0.001	-0.0003	0.0002	414,872	2.25	2511	2795	1042	-462	541,727	3.24
5	-460	-543	-403	145	364,469	2.18	-1922	-2181	-809	344	300,364	2.50
9	3392	3706	1421	-685	83,600	2.16	4024	5176	1581	-779	84,242	2.42
	80% Scenario											
1	4639	5217	1813	-1037	543,488	3.61	-4753	-5446	-1848	1115	390,700	4.01
2	-6880	-7702	-2649	1658	141,377	3.73	378	917	63	-117	181,275	3.84
4	2.09	2.45	0.79	-0.59	533,774	2.06	-2316	-2570	-819	547	650,472	3.04
5	133	334	284	-110	449,143	2.11	3589	4001	1382	-801.8	366,463	2.24
9	5430	6137	2205	-1330	106,827	2.10	2338	2662	917	-540.2	98,066	2.29

Notes:

1. No other control variables are included in the model.
2. Occupations with employment falling in the upper and lower five percentile are dropped after adjustment.
3. The maximum iterations is500 times.
4. Initial value is adopted from the first-round cubic spline interpolation results without dropping any observations. See Table 3.A for details of initial value.

Table 3.7.1: Calculated Number of Jobs Offshored and Offshoring Percentage for Group 1 (Management, Business and Financial Operations Occupations) from NLS Method

	Year	20% Scenario		40% Scenario		80% Scenario	
		No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage
Pre-2000	1983	7,771,465	47.6%	10,400,000	55.1%	13,200,000	61.1%
	1984	7,085,993	45.8%	9,713,385	53.5%	12,500,000	59.7%
	1985	6,707,947	46.1%	9,335,339	53.8%	12,200,000	60.0%
	1986	6,292,123	42.1%	8,919,514	50.3%	11,700,000	57.0%
	1987	5,989,792	40.0%	8,617,183	48.5%	11,400,000	55.4%
	1988	5,500,926	37.7%	8,128,317	46.6%	11,000,000	53.7%
	1989	4,891,834	35.6%	7,519,225	44.8%	10,300,000	52.2%
	1990	4,902,728	34.8%	7,530,119	44.1%	10,400,000	51.6%
	1991	4,671,854	34.7%	7,299,245	44.0%	10,100,000	51.5%
	1992	6,299,921	36.2%	8,927,313	45.4%	11,800,000	52.6%
	1993	6,233,880	32.9%	8,861,271	42.5%	11,700,000	50.2%
	1994	6,354,836	35.1%	8,982,227	44.3%	11,800,000	51.8%
	1995	5,822,961	33.1%	8,450,352	42.6%	11,300,000	50.3%
	1996	5,651,819	33.3%	8,279,210	42.9%	11,100,000	50.5%
	1997	5,185,593	32.9%	7,812,984	42.5%	10,600,000	50.2%
	1998	4,869,771	30.6%	7,497,162	40.5%	10,300,000	48.4%
	1999	4,454,344	27.5%	7,081,736	37.9%	9,911,625	46.2%
Post-2000	2000	7,696,804	37.5%	11,100,000	46.4%	16,000,000	55.6%
	2001	7,312,754	35.2%	10,700,000	44.4%	15,600,000	54.0%
	2002	7,152,252	36.1%	10,500,000	45.2%	15,500,000	54.7%
	2003	7,336,184	37.0%	10,700,000	46.0%	15,700,000	55.3%
	2004	7,286,110	35.7%	10,700,000	44.9%	15,600,000	54.4%
	2005	7,108,696	34.0%	10,500,000	43.4%	15,400,000	53.2%
	2006	6,919,571	34.1%	10,300,000	43.5%	15,300,000	53.2%
	2007	6,652,990	33.0%	10,000,000	42.6%	15,000,000	52.5%
	2008	6,526,403	34.0%	9,920,160	43.4%	14,900,000	53.1%
	2009	6,630,334	32.5%	10,000,000	42.2%	15,000,000	52.1%
	2010	6,953,488	32.1%	10,300,000	41.8%	15,300,000	51.9%
	2011	6,892,056	33.6%	10,300,000	43.1%	15,200,000	52.9%

Notes: 1. The number of job offshored is the sum of job offshored across all occupations within an occupational group.
2. The offshoring percentage is the average offshoring percentage across all occupations within an occupational group.
3. A cubic offshoring cost function is assumed.

Table 3.7.2: Calculated Number of Jobs Offshored and Offshoring Percentage for Group 2 (Professional and Related Occupations) from NLS Method

	Year	20% Scenario		40% Scenario		80% Scenario	
		No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage
Pre-2000	1983	13,300,000	46.8%	17,800,000	54.4%	23,400,000	61.2%
	1984	13,200,000	46.6%	17,700,000	54.3%	23,300,000	61.1%
	1985	12,800,000	46.1%	17,300,000	53.8%	22,900,000	60.7%
	1986	12,500,000	44.8%	17,100,000	52.7%	22,600,000	59.8%
	1987	12,400,000	45.5%	16,900,000	53.3%	22,500,000	60.3%
	1988	12,000,000	43.1%	16,500,000	51.3%	22,100,000	58.5%
	1989	11,700,000	44.0%	16,300,000	52.0%	21,800,000	59.2%
	1990	11,200,000	40.9%	15,700,000	49.4%	21,300,000	56.9%
	1991	11,100,000	39.5%	15,700,000	48.2%	21,200,000	55.9%
	1992	11,000,000	41.0%	15,500,000	49.4%	21,100,000	57.0%
	1993	10,600,000	39.1%	15,100,000	47.8%	20,700,000	55.6%
	1994	10,900,000	40.1%	15,500,000	48.6%	21,000,000	56.3%
	1995	10,600,000	40.1%	15,100,000	48.7%	20,700,000	56.4%
	1996	10,100,000	38.9%	14,700,000	47.6%	20200,000	55.4%
	1997	9,780,010	37.5%	14,300,000	46.4%	19,900,000	54.4%
	1998	9,640,700	36.4%	14,200,000	45.5%	19,700,000	53.6%
	1999	8,816,018	34.8%	13,300,000	44.1%	18,900,000	52.5%
Post-2000	2000	12,300,000	39.4%	17,300,000	48.1%	24,800,000	57.2%
	2001	11,800,000	38.3%	16,800,000	47.1%	24,200,000	56.4%
	2002	11,600,000	38.5%	16,600,000	47.3%	24,000,000	56.5%
	2003	11,500,000	37.8%	16,500,000	46.7%	24,000,000	56.1%
	2004	11,400,000	37.4%	16,400,000	46.3%	23,800,000	55.7%
	2005	11,200,000	37.2%	16,200,000	46.1%	23,600,000	55.6%
	2006	11,000,000	36.1%	16,000,000	45.2%	23,500,000	54.8%
	2007	10,400,000	35.5%	15,400,000	44.7%	22,900,000	54.4%
	2008	10,100,000	35.0%	15,100,000	44.3%	22,500,000	54.1%
	2009	9,978,409	34.9%	15,000,000	44.2%	22,400,000	54.0%
	2010	10,000,000	34.2%	15,000,000	43.6%	22,500,000	53.5%
	2011	9,962,497	34.2%	15,000,000	43.6%	22,400,000	53.5%

Notes: 1. The number of job offshored is the sum of job offshored across all occupations within an occupational group.
2. The offshoring percentage is the average offshoring percentage across all occupations within an occupational group.
3. A cubic offshoring cost function is assumed.

Table 3.7.3: Calculated Number of Jobs Offshored and Offshoring Percentage for Group 4 (Sales and Related Occupations) from NLS Method

	Year	20% Scenario		40% Scenario		80% Scenario	
		No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage
Pre-2000	1983	4,489,843	35.1%	6,585,252	44.3%	10,800,000	56.7%
	1984	4,182,837	32.5%	6,278,245	42.1%	10,500,000	55.0%
	1985	4,178,781	32.4%	6,274,190	42.0%	10,500,000	54.9%
	1986	4,014,622	30.7%	6,110,031	40.6%	10,300,000	53.8%
	1987	3,805,164	29.2%	5,900,573	39.3%	10,100,000	52.8%
	1988	3,846,256	30.6%	5,941,665	40.5%	10,200,000	53.7%
	1989	3,670,658	30.7%	5,766,067	40.6%	9,987,918	53.8%
	1990	3,370,897	29.0%	5,466,306	39.1%	9,688,157	52.6%
	1991	3,407,134	30.8%	5,502,543	40.7%	9,724,394	53.9%
	1992	3,370,818	27.4%	5,466,226	37.8%	9,688,077	51.6%
	1993	3,440,809	27.9%	5,536,218	38.2%	9,758,068	51.9%
	1994	4,456,491	34.7%	6,551,899	44.0%	10,800,000	56.5%
	1995	4,284,704	32.7%	6,380,113	42.3%	10,600,000	55.1%
	1996	4,027,732	31.4%	6,123,140	41.1%	10,300,000	54.3%
	1997	3,933,802	31.0%	6,029,211	40.8%	10,300,000	54.0%
	1998	3,888,172	30.2%	5,983,581	40.2%	10,200,000	53.5%
	1999	4,005,604	28.9%	6,101,013	39.0%	10,300,000	52.6%
Post-2000	2000	5,993,041	32.1%	9,110,449	41.8%	5,993,041	32.1%
	2001	5,701,304	28.8%	8,818,713	39.0%	5,701,304	28.8%
	2002	5,635,168	28.9%	8,752,576	39.1%	5,635,168	28.9%
	2003	5,805,051	29.9%	8,922,459	39.9%	5,805,051	29.9%
	2004	5,677,808	29.7%	8,795,216	39.7%	5,677,808	29.7%
	2005	5,556,994	30.0%	8,674,402	40.0%	5,556,994	30.0%
	2006	5,363,608	27.9%	8,481,016	38.2%	5,363,608	27.9%
	2007	5,251,183	28.0%	8,368,592	38.3%	5,251,183	28.0%
	2008	5480,845	29.9%	8,598,254	39.9%	5,480,845	29.9%
	2009	5,400,839	29.6%	8,518,247	39.7%	5,400,839	29.6%
	2010	5,567,462	31.5%	8,684,871	41.3%	5,567,462	31.5%
	2011	5,655,523	32.4%	8,772,932	42.0%	5,655,523	32.4%

Notes: 1. The number of job offshored is the sum of job offshored across all occupations within an occupational group.
2. The offshoring percentage is the average offshoring percentage across all occupations within an occupational group.
3. A cubic offshoring cost function is assumed.

Table 3.7.4: Calculated Number of Jobs Offshored and Offshoring Percentage for Group 5 (Office and Administrative Support Occupations) from NLS Method

	Year	20% Scenario		40% Scenario		80% Scenario	
		No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage
Pre-2000	1983	10,200,000	41.9%	14,600,000	50.2%	21,800,000	59.6%
	1984	10,400,000	43.7%	14,800,000	51.7%	21,900,000	60.8%
	1985	10,200,000	43.1%	14,600,000	51.2%	21,700,000	60.4%
	1986	10,100,000	41.8%	14,500,000	50.1%	21,600,000	59.5%
	1987	9,897,058	40.4%	14,300,000	48.9%	21,400,000	58.5%
	1988	9,852,361	41.0%	14,200,000	49.5%	21,400,000	59.0%
	1989	9,901,249	39.6%	14,300,000	48.2%	21,400,000	58.0%
	1990	9,387,039	36.5%	13,800,000	45.6%	20,900,000	55.8%
	1991	9,582,273	36.2%	14,000,000	45.3%	21,100,000	55.6%
	1992	9,565,003	36.1%	13,900,000	45.2%	21,100,000	55.5%
	1993	9,722,870	35.8%	14,100,000	44.9%	21,200,000	55.3%
	1994	11,800,000	42.2%	16,200,000	50.5%	23,400,000	59.8%
	1995	11,800,000	42.5%	16,200,000	50.7%	23,300,000	60.0%
	1996	11,700,000	41.6%	16,100,000	49.9%	23,200,000	59.4%
	1997	11,800,000	40.4%	16,200,000	48.9%	23,300,000	58.5%
	1998	11,800,000	39.5%	16,200,000	48.2%	23,300,000	57.9%
	1999	11,800,000	40.3%	16,200,000	48.8%	23,300,000	58.5%
Post-2000	2000	10,600,000	36.7%	15,400,000	45.7%	22,900,000	55.5%
	2001	10,600,000	38.6%	15,400,000	47.4%	23,000,000	56.9%
	2002	10,900,000	39.9%	15,800,000	48.5%	23,300,000	57.8%
	2003	11,000,000	36.6%	15,800,000	45.7%	23,300,000	55.5%
	2004	11,100,000	36.1%	16,000,000	45.2%	23,500,000	55.1%
	2005	11,100,000	35.3%	15,900,000	44.5%	23,400,000	54.5%
	2006	11,000,000	36.6%	15,900,000	45.6%	23,400,000	55.4%
	2007	11,200,000	38.8%	16,100,000	47.6%	23,600,000	57.0%
	2008	11,000,000	38.3%	15,900,000	47.1%	23,400,000	56.7%
	2009	11,500,000	39.9%	16,400,000	48.5%	23,900,000	57.8%
	2010	11,600,000	39.3%	16,500,000	48.0%	24,000,000	57.3%
	2011	12,000,000	40.3%	16,900,000	48.9%	24,400,000	58.1%

Notes: 1.The number of job offshored is the sum of job offshored across all occupations within an occupational group.
2. The offshoring percentage is the average offshoring percentage across all occupations within an occupational group.
3. A cubic offshoring cost function is assumed.

Table 3.7.5: Calculated Number of Jobs Offshored and Offshoring Percentage for Group 9 (Production Occupations) from NLS Method

	Year	20% Scenario		40% Scenario		80% Scenario	
		No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage	No. of Jobs Offshored	Offshore Percentage
Pre-2000	1983	6,709,289	36.5%	9,768,267	45.6%	15,700,000	57.4%
	1984	6,856,422	38.4%	9,915,399	47.2%	15,800,000	58.7%
	1985	6,862,317	37.9%	9,921,295	46.8%	15,900,000	58.4%
	1986	6,864,584	37.9%	9,923,561	46.8%	15,900,000	58.4%
	1987	6,944,952	38.7%	10,000,000	47.5%	15,900,000	58.9%
	1988	7,105,089	38.4%	10,200,000	47.2%	16,100,000	58.7%
	1989	6,948,416	37.9%	10,000,000	46.8%	15,900,000	58.4%
	1990	6,878,102	37.1%	9,937,079	46.1%	15,900,000	57.8%
	1991	7,062,489	38.8%	10,100,000	47.6%	16,100,000	59.0%
	1992	7,208,100	39.1%	10,300,000	47.8%	16,200,000	59.2%
	1993	7,240,489	40.0%	10,300,000	48.6%	16,200,000	59.8%
	1994	8,090,008	43.1%	11,100,000	51.2%	17,100,000	61.8%
	1995	8,044,484	43.7%	11,100,000	51.8%	17,000,000	62.3%
	1996	7,987,601	44.3%	11,000,000	52.3%	17,000,000	62.7%
	1997	8,040,568	44.1%	11,100,000	52.1%	17,000,000	62.5%
	1998	8,274,450	44.4%	11,300,000	52.3%	17,300,000	62.7%
	1999	8,496,861	46.3%	11,600,000	54.0%	17,500,000	64.0%
Post-2000	2000	4,711,486	36.1%	6,992,078	45.2%	9,611,898	52.9%
	2001	4,922,465	36.6%	7,203,058	45.6%	9,822,878	53.3%
	2002	5,399,961	40.9%	7,680,554	49.4%	10,300,000	56.5%
	2003	5,928,591	42.0%	8,209,184	50.3%	10,800,000	57.3%
	2004	6,092,295	44.3%	8,372,887	52.2%	11,000,000	59.0%
	2005	6,095,895	44.4%	8,376,488	52.4%	11,000,000	59.1%
	2006	6,241,160	45.6%	8,521,752	53.4%	11,100,000	60.0%
	2007	6,225,142	45.7%	8,505,735	53.5%	11,100,000	60.0%
	2008	6,290,275	47.3%	8,570,868	54.8%	11,200,000	61.2%
	2009	6,905,819	49.8%	9,186,411	57.0%	11,800,000	63.0%
	2010	6,962,225	49.1%	9,242,818	56.4%	11,900,000	62.5%
	2011	6,812,466	48.5%	9,093,059	55.9%	11,700,000	62.1%

Notes: 1. The number of job offshored is the sum of job offshored across all occupations within an occupational group.
2. The offshoring percentage is the average offshoring percentage across all occupations within an occupational group.
3. A cubic offshoring cost function is assumed.

Table 3.8: Scenario Comparison among the Five Relatively Offshorable Occupational Groups from NLS Method

	Year	Group 1 (20%)		Group 2 (20%)		Group 4 (20%)		Group 5 (40%)		Group 9 (80%)	
		Offshored Jobs	%	Offshored Jobs	%	Offshored Jobs	%	Offshored Jobs	%	Offshored Jobs	%
Pre-2000	1983	7,771,465	47.6%	13,300,000	46.8%	4,489,843	35.1%	14,600,000	50.2%	15,700,000	57.4%
	1984	7,085,993	45.8%	13,200,000	46.6%	4,182,837	32.5%	14,800,000	51.7%	15,800,000	58.7%
	1985	6,707,947	46.1%	12,800,000	46.1%	4,178,781	32.4%	14,600,000	51.2%	15,900,000	58.4%
	1986	6,292,123	42.1%	12,500,000	44.8%	4,014,622	30.7%	14,500,000	50.1%	15,900,000	58.4%
	1987	5,989,792	40.0%	12,400,000	45.5%	3,805,164	29.2%	14,300,000	48.9%	15,900,000	58.9%
	1988	5,500,926	37.7%	12,000,000	43.1%	3,846,256	30.6%	14,200,000	49.5%	16,100,000	58.7%
	1989	4,891,834	35.6%	11,700,000	44.0%	3,670,658	30.7%	14,300,000	48.2%	15,900,000	58.4%
	1990	4,902,728	34.8%	11,200,000	40.9%	3,370,897	29.0%	13,800,000	45.6%	15,900,000	57.8%
	1991	4,671,854	34.7%	11,100,000	39.5%	3,407,134	30.8%	14,000,000	45.3%	16,100,000	59.0%
	1992	6,299,921	36.2%	11,000,000	41.0%	3,370,818	27.4%	13,900,000	45.2%	16,200,000	59.2%
	1993	6,233,880	32.9%	10,600,000	39.1%	3,440,809	27.9%	14,100,000	44.9%	16,200,000	59.8%
	1994	6,354,836	35.1%	10,900,000	40.1%	4,456,491	34.7%	16,200,000	50.5%	17,100,000	61.8%
	1995	5,822,961	33.1%	10,600,000	40.1%	4,284,704	32.7%	16,200,000	50.7%	17,000,000	62.3%
	1996	5,651,819	33.3%	10,100,000	38.9%	4,027,732	31.4%	16,100,000	49.9%	17,000,000	62.7%
	1997	5,185,593	32.9%	9,780,010	37.5%	3,933,802	31.0%	16,200,000	48.9%	17,000,000	62.5%
	1998	4,869,771	30.6%	9,640,700	36.4%	3,888,172	30.2%	16,200,000	48.2%	17,300,000	62.7%
	1999	4,454,344	27.5%	8,816,018	34.8%	4,005,604	28.9%	16,200,000	48.8%	17,500,000	64.0%

Table 3.8 (cont'd)

Post-2000	2000	7,696,804	37.5%	12,300,000	39.4%	5,993,041	32.1%	15,400,000	45.7%	9,611,898	52.9%
	2001	7,312,754	35.2%	11,800,000	38.3%	5,701,304	28.8%	15,400,000	47.4%	9,822,878	53.3%
	2002	7,152,252	36.1%	11,600,000	38.5%	5,635,168	28.9%	15,800,000	48.5%	10,300,000	56.5%
	2003	7,336,184	37.0%	11,500,000	37.8%	5,805,051	29.9%	15,800,000	45.7%	10,800,000	57.3%
	2004	7,286,110	35.7%	11,400,000	37.4%	5,677,808	29.7%	16,000,000	45.2%	11,000,000	59.0%
	2005	7,108,696	34.0%	11,200,000	37.2%	5,556,994	30.0%	15,900,000	44.5%	11,000,000	59.1%
	2006	6,919,571	34.1%	11,000,000	36.1%	5,363,608	27.9%	15,900,000	45.6%	11,100,000	60.0%
	2007	6,652,990	33.0%	10,400,000	35.5%	5,251,183	28.0%	16,100,000	47.6%	11,100,000	60.0%
	2008	6,526,403	34.0%	10,100,000	35.0%	5480,845	29.9%	15,900,000	47.1%	11,200,000	61.2%
	2009	6,630,334	32.5%	9,978,409	34.9%	5,400,839	29.6%	16,400,000	48.5%	11,800,000	63.0%
	2010	6,953,488	32.1%	10,000,000	34.2%	5,567,462	31.5%	16,500,000	48.0%	11,900,000	62.5%
	2011	6,892,056	33.6%	9,962,497	34.2%	5,655,523	32.4%	16,900,000	48.9%	11,700,000	62.1%

Notes: 1. The number of job offshored is the sum of job offshored across all occupations within an occupational group.
2. The offshoring percentage is the average offshoring percentage across all occupations within an occupational group.
3. A cubic offshoring cost function is assumed.
4. In 20%, 40% and 80% scenario, offshoring is set not to exceed the 20%, 40% and 80% of the maximum adjusted employment of all occupations across all years within each occupational group respectively.

Table 3.A: Monotonic Cubic Spline Interpolation Method Preliminary Point Estimates of Parameterized Offshoring Cost Function $t(i)$

Group	Value of $t(i)$ at Control Points i Pre-2000: 1983-1999										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	3.02	8.98	14.1	26.2	36.3	48.6	61.7	80.1	98.83	126.5	185.1
2	1.11	2.65	8.75	15.4	32.5	72.0	121.4	138.1	131.7	123.4	109.1
3	8.40	6.90	12.8	18.7	47.3	174.3	262.4	372.7	298.5	452.6	259.2
4	0.04	7.00	12.3	20.6	38.8	57.8	87.8	120.9	185.8	167.1	144.0
5	13.9	5.53	11.5	20.2	30.6	43.2	59.5	79.1	103.7	131.8	146.1
6	8.82	13.4	20.2	32.6	46.2	83.6	119.4	182.9	272.1	401.7	537.2
7	5.37	9.24	17.2	25.2	52.4	75.5	92.5	136.1	191.1	307.4	438.8
8	8.09	0.12	15.3	14.7	75.6	110.6	131.8	172.0	266.1	355.9	552.8
9	8.29	0.14	0.94	7.27	18.9	54.0	104.4	136.7	151.5	173.3	165.1
10	8.25	10.1	15.8	35.6	71.7	208.7	284.9	421.4	561.2	798.0	1011.5
	Value of $t(i)$ at Control Points i Post-2000 Period: 2000-2011										
1	3.09	6.85	12.6	20.5	30.8	43.05	58.9	78.18	102.2	121.8	186.0
2	8.49	2.09	4.45	7.80	29.7	110.2	157.6	197.4	225.4	173.1	459.4
3	17.4	26.9	55.4	69.5	141.6	232.7	180.7	405.3	533.4	744.8	1076.7
4	0.87	6.92	9.56	19.8	24.4	51.8	59.7	86.4	102.0	129.5	184.9
5	3.27	1.01	0.88	37.2	33.8	48.0	64.3	85.5	101.7	133.3	165.2
6	22.4	23.2	25.3	35.7	53.5	77.0	111.3	165.6	249.8	215.3	259.1
7	2.21	27.9	26.6	40.8	23.2	82.3	116.8	203.8	265.3	369.0	533.2
8	5.69	7.30	2.14	18.2	34.6	41.3	62.5	92.1	134.7	182.3	269.8
9	0.70	0.09	1.32	40.6	19.5	71.3	105.2	95.2	85.91	117.7	124.6
10	16.1	11.2	11.4	53.7	79.8	99.4	132.5	171.3	124.0	195.0	241.0

Notes: 1. No other control variables are included in the model.

2. The upper bound is set that offshoring cannot exceed the 10% of the maximum employment of all occupations across all years within each group.

3. No observations are dropped.

4. The maximum iterations is500 times.

5. For Group 1, 2, 4, 5, 9the functional form for iteration to start with is $(4x + 1.5)^3 + \varepsilon$, where ε is a random shock with normal distribution $N(0,\ 0.01)$.

For Group 3, 6, 7, 8 and 10, the functional form for iteration to start with is $10 * exp(4x) + \varepsilon$, where ε is a random shock with normal distribution $N(0,\ 0.01)$.

Figure 3.2.1: Monotonic Cubic Spline Interpolation Method Offshoring Cost Function by Occupational Groups in Pre-2000 Period (1983-1999)

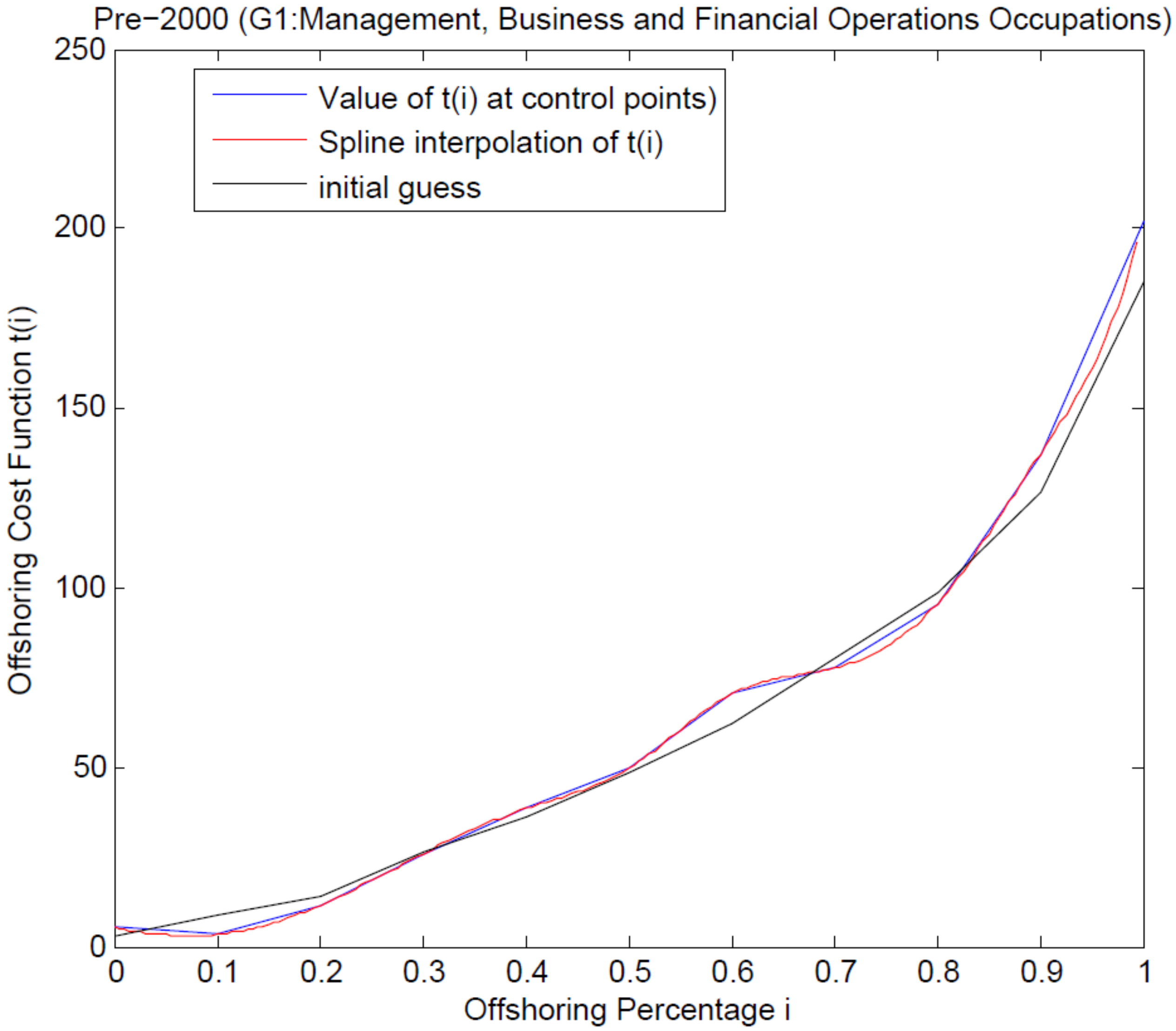


Figure 3.2.1 (cont'd)

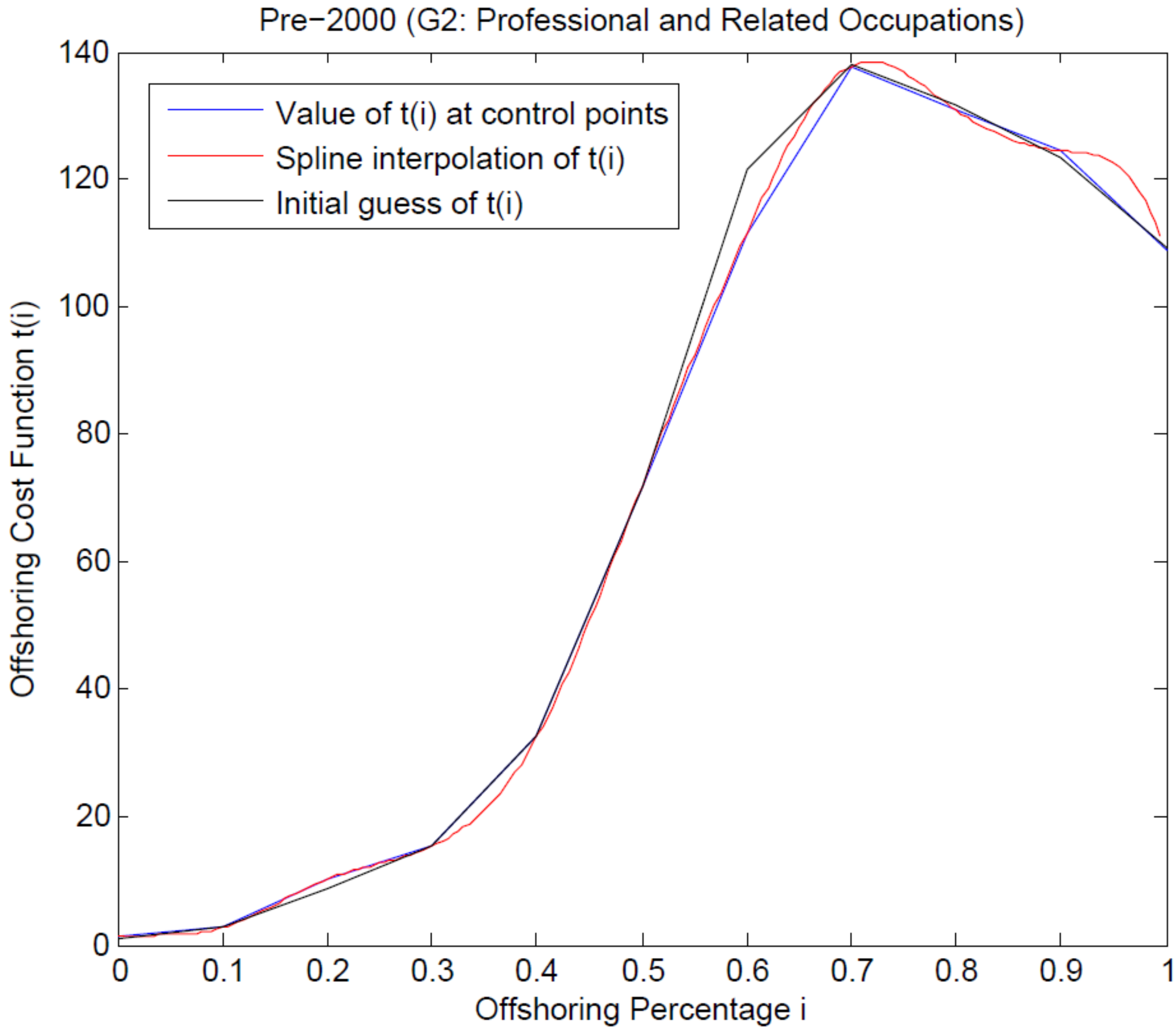


Figure 3.2.1 (cont'd)

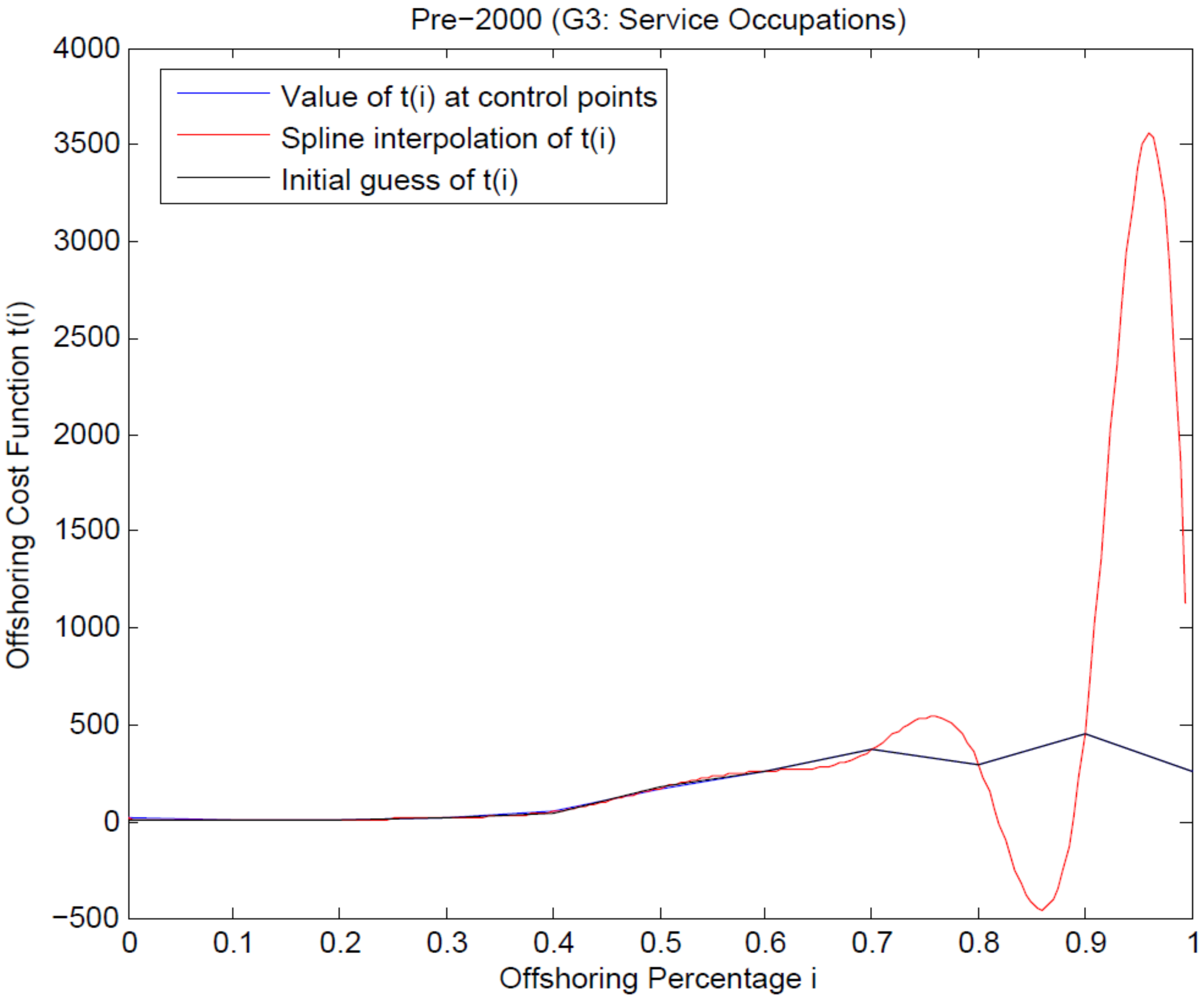


Figure 3.2.1 (cont'd)

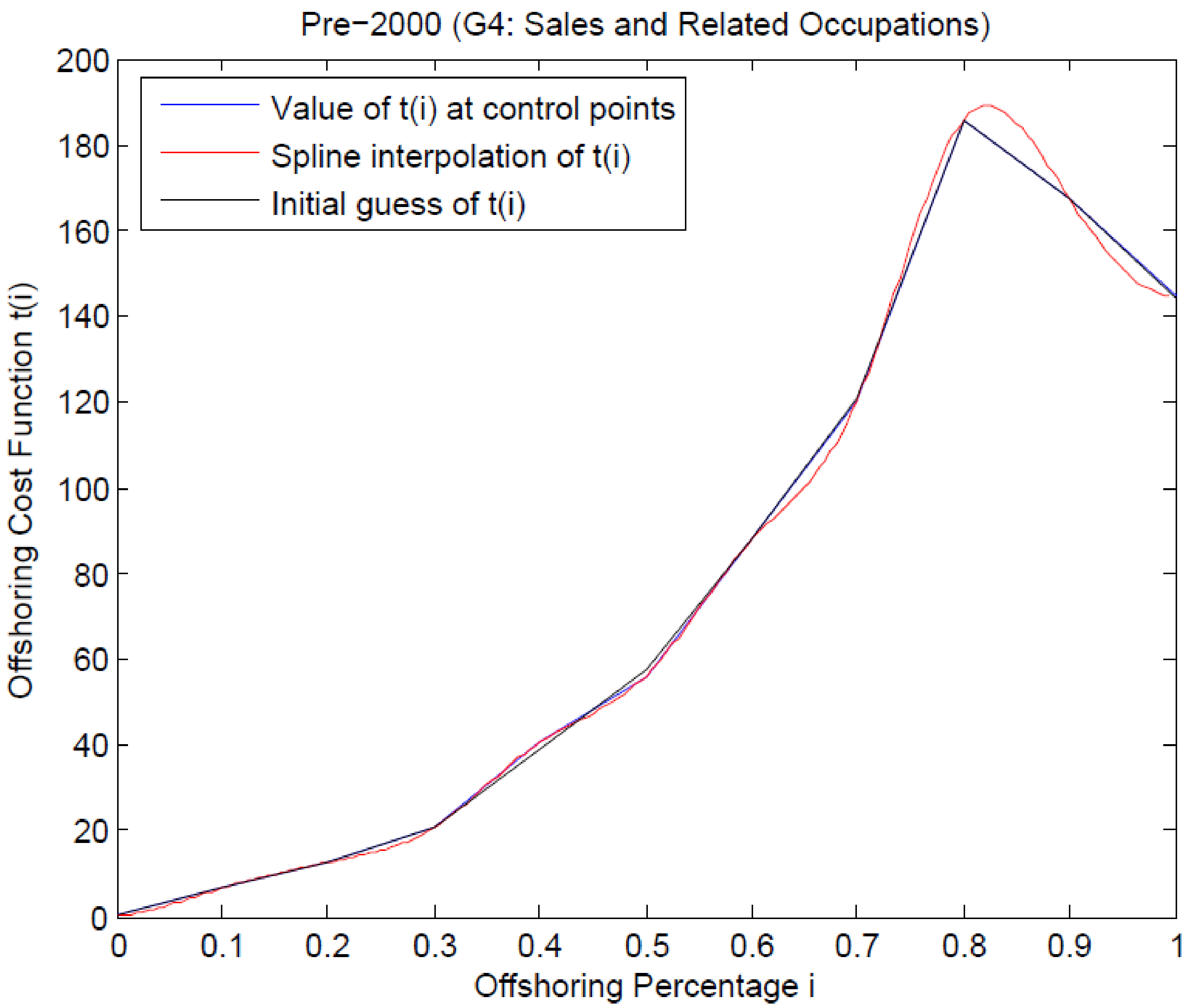


Figure 3.2.1 (cont'd)

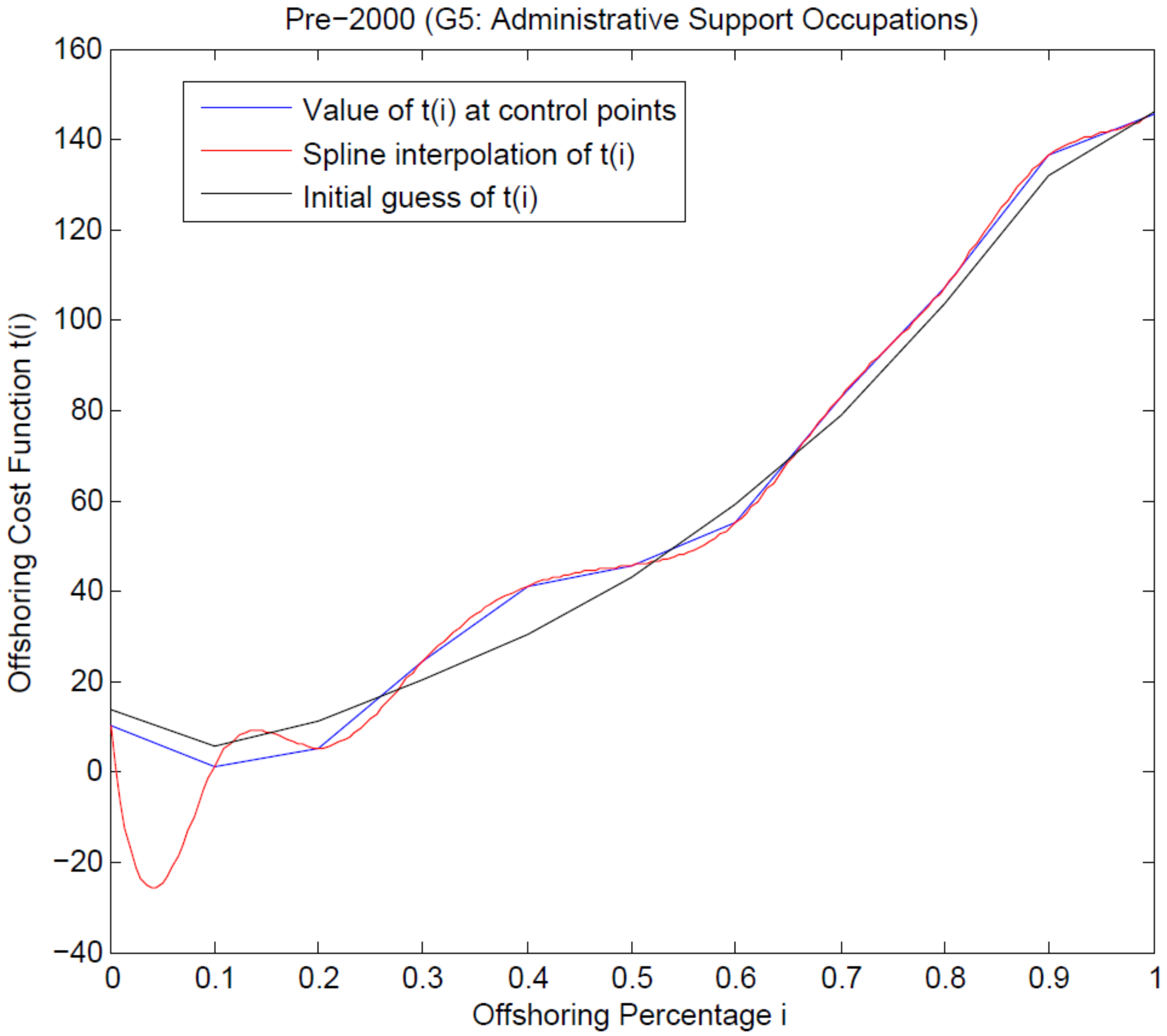


Figure 3.2.1 (cont'd)

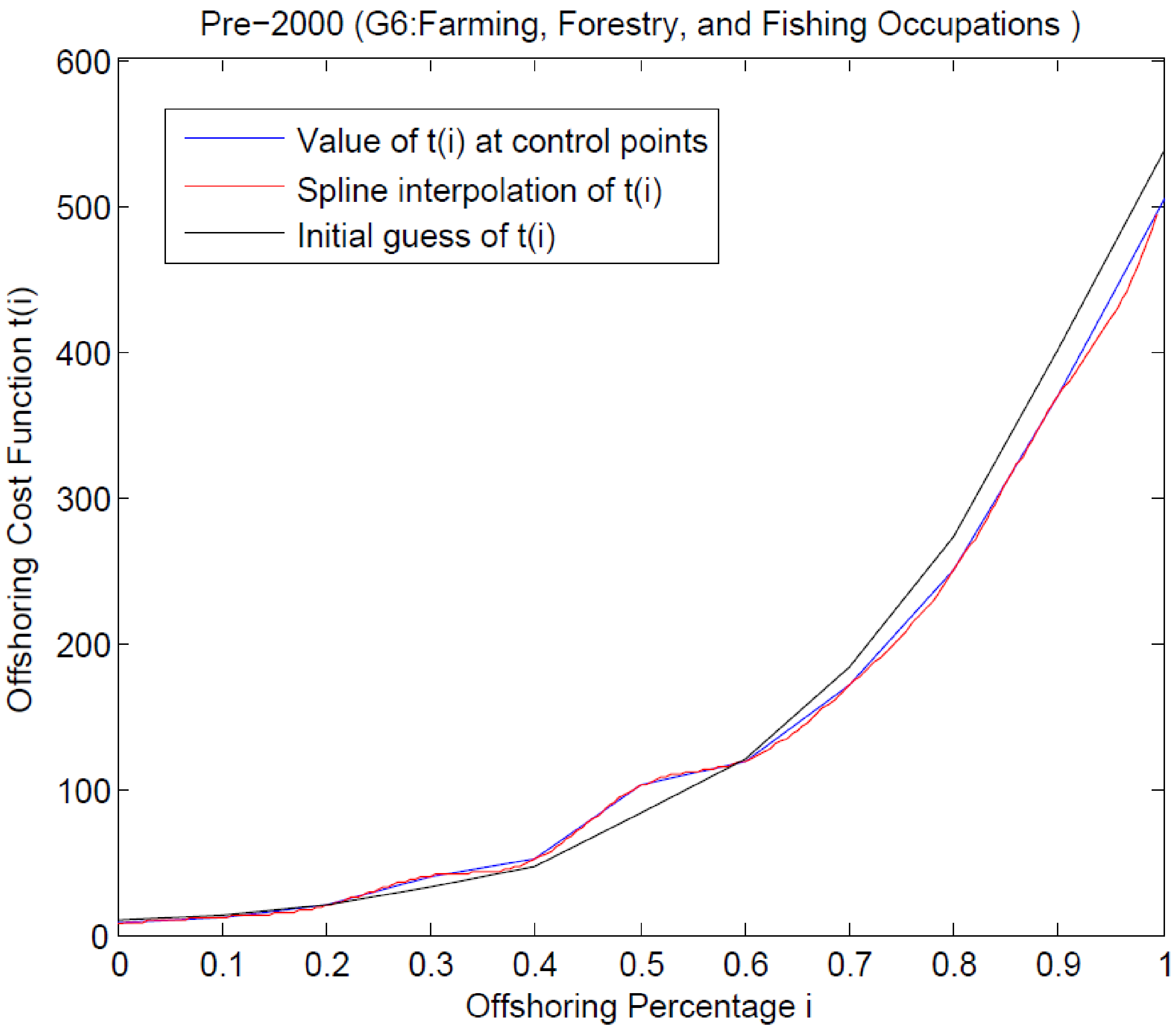


Figure 3.2.1 (cont'd)

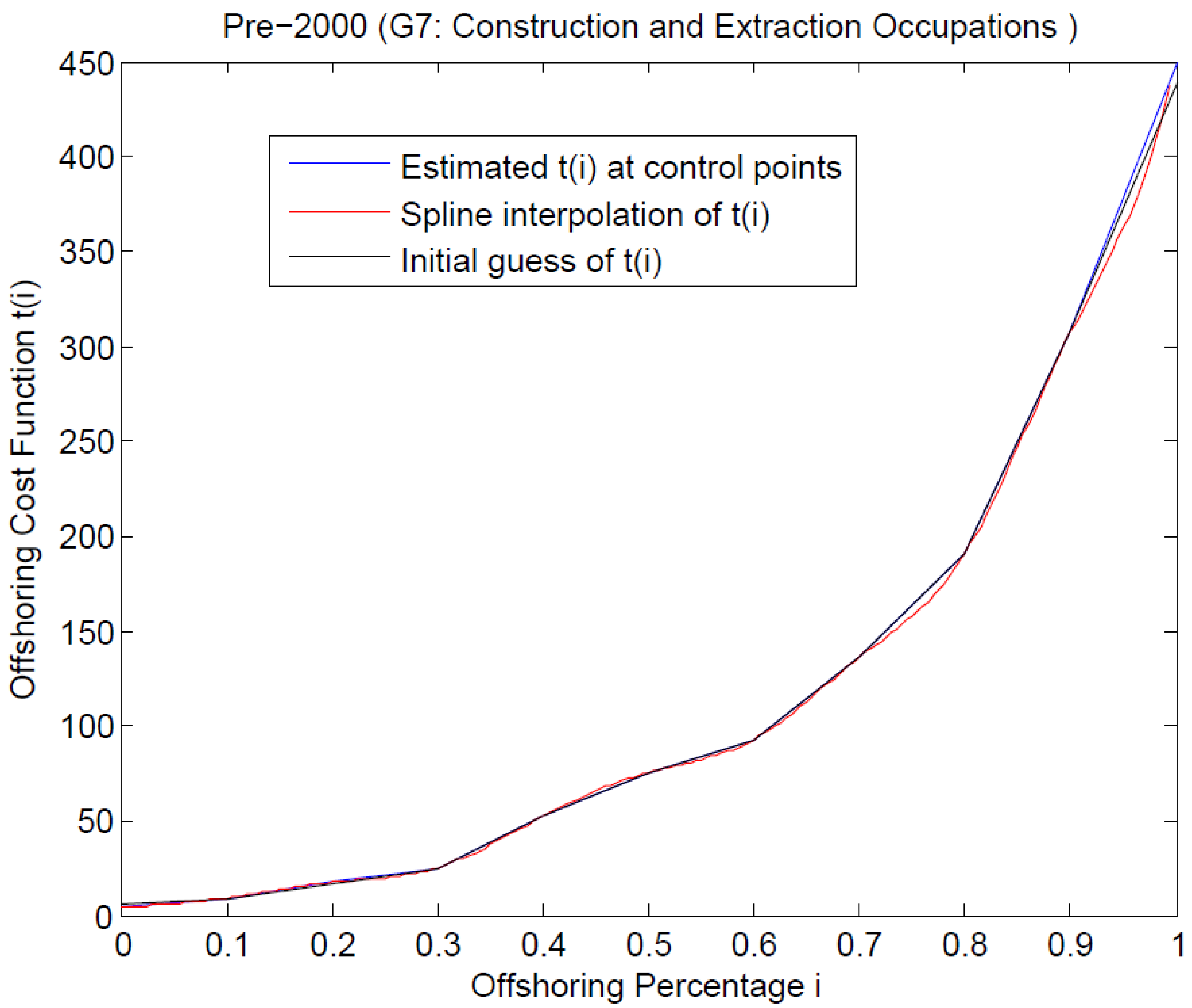


Figure 3.2.1 (cont'd)

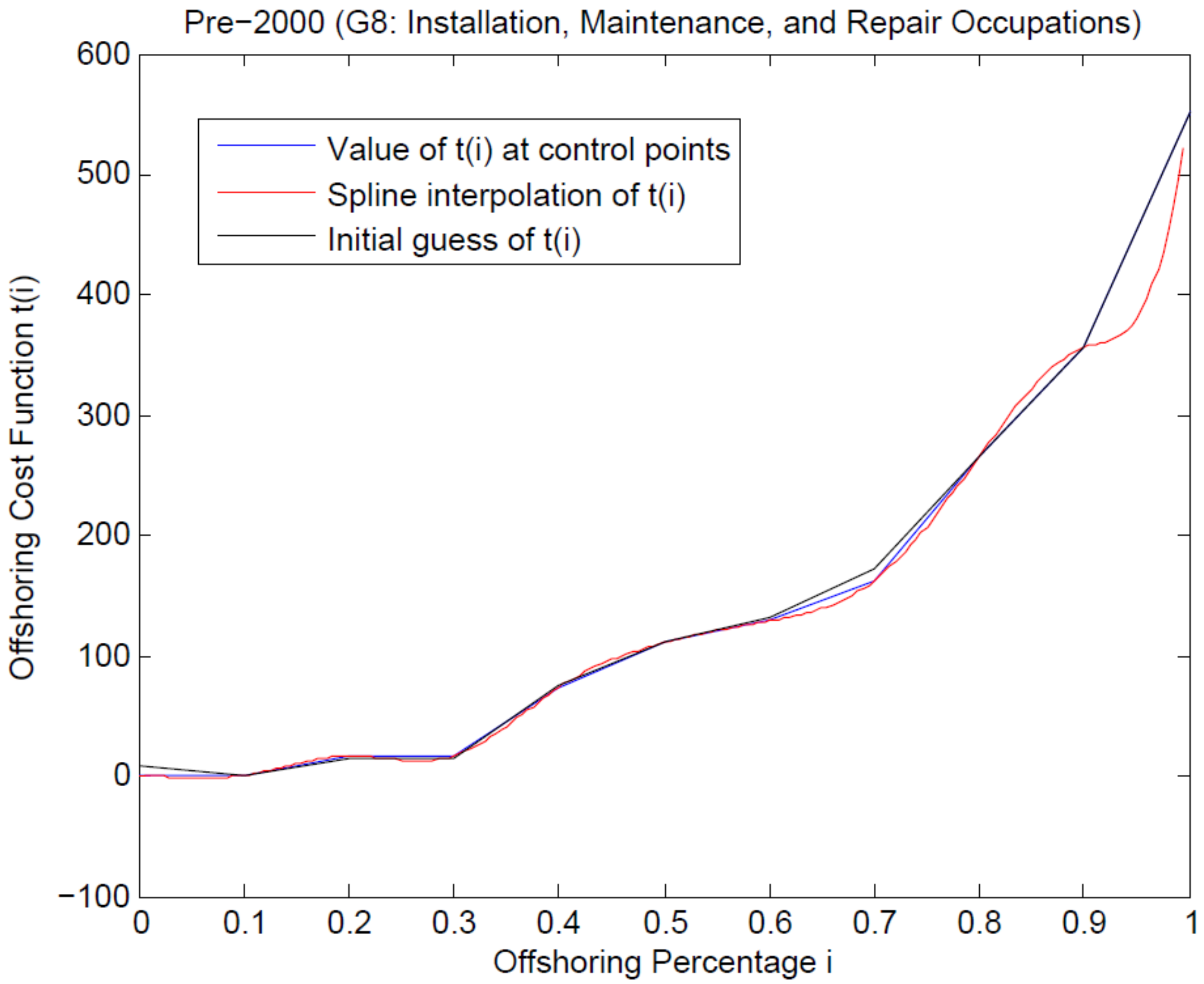


Figure 3.2.1 (cont'd)

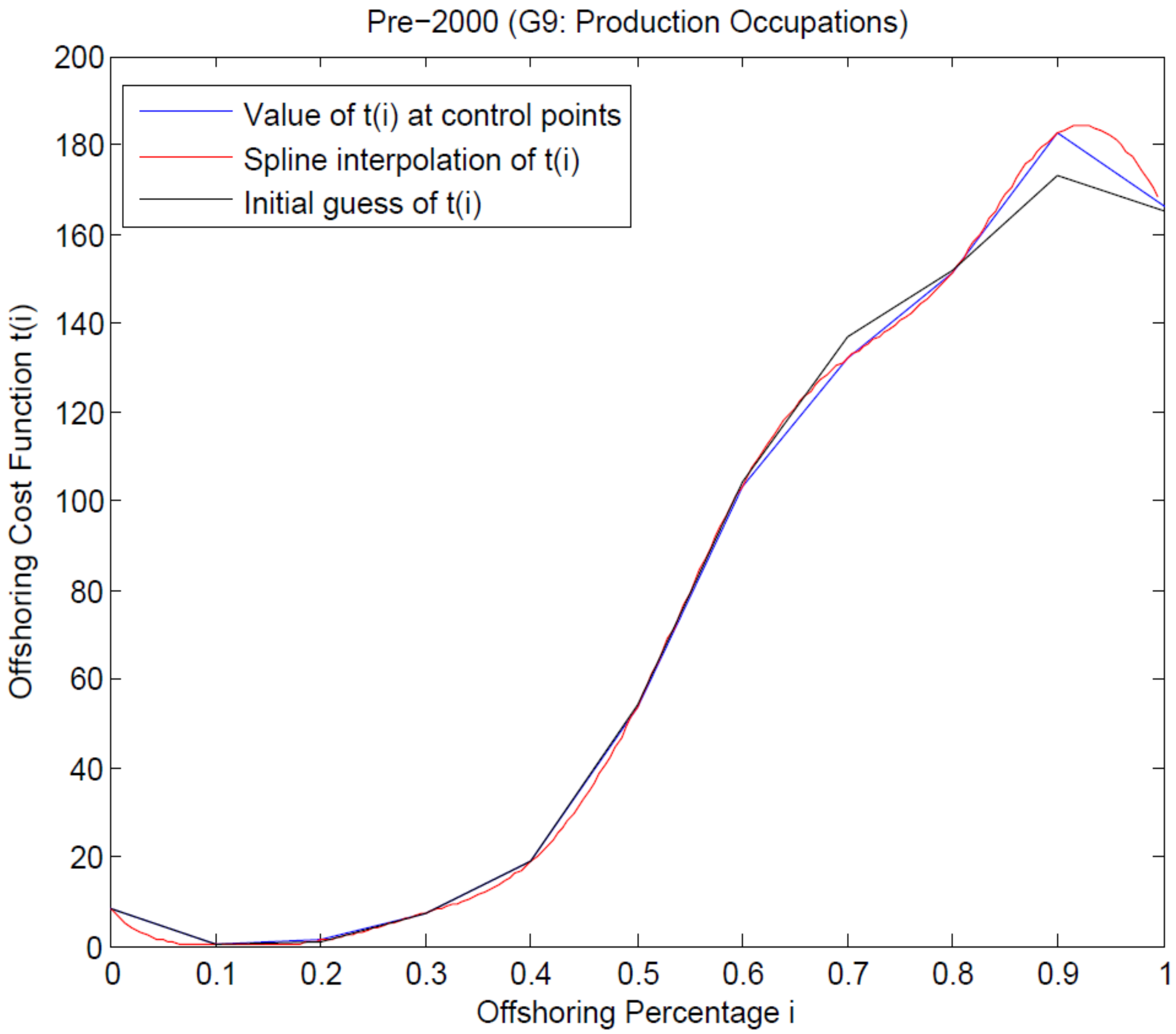


Figure 3.2.1 (cont'd)

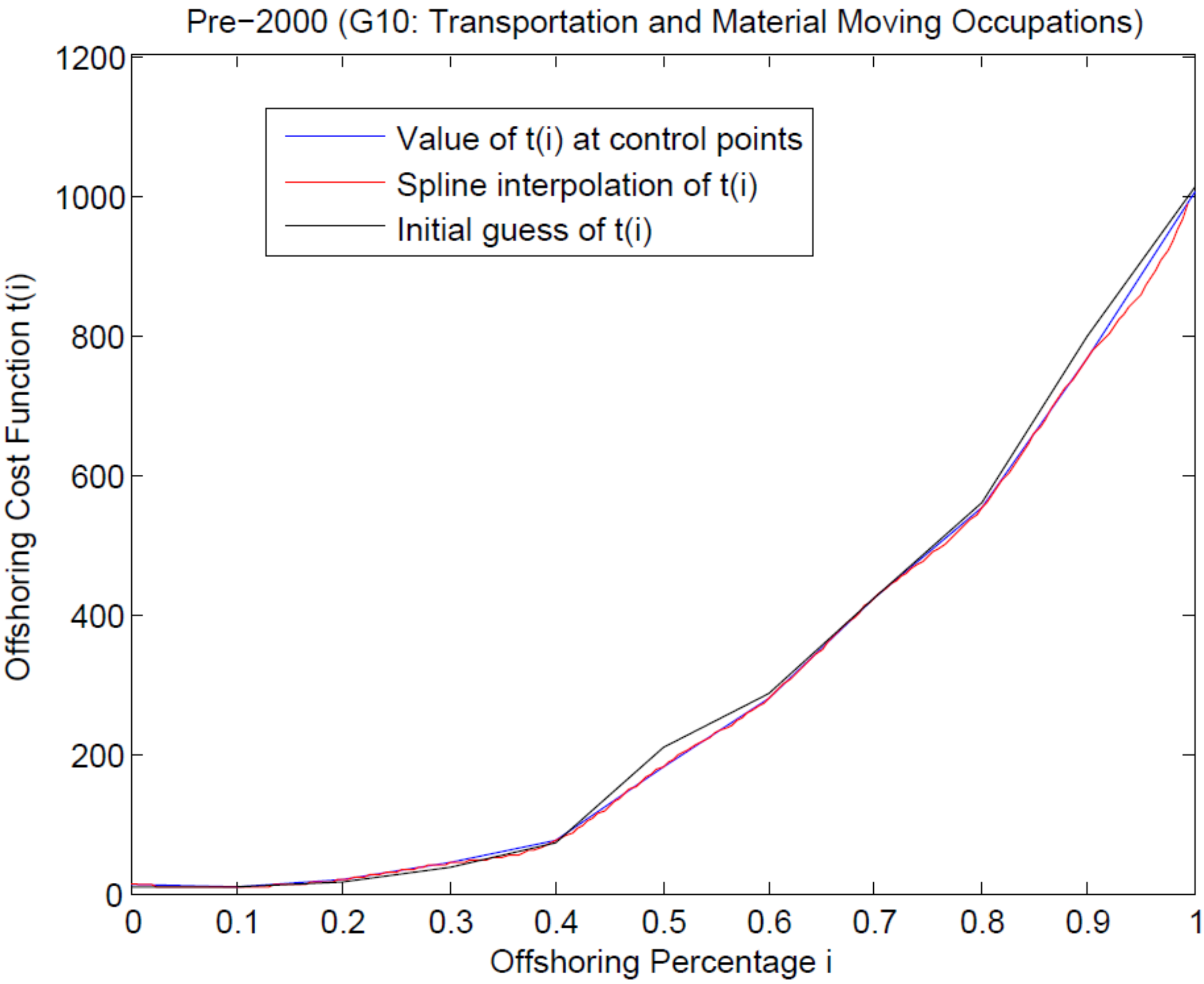


Figure 3.2.2: Monotonic Cubic Spline Interpolation Method Offshoring Cost Function by Occupational Groups in Post-2000 Period (2000-2011)

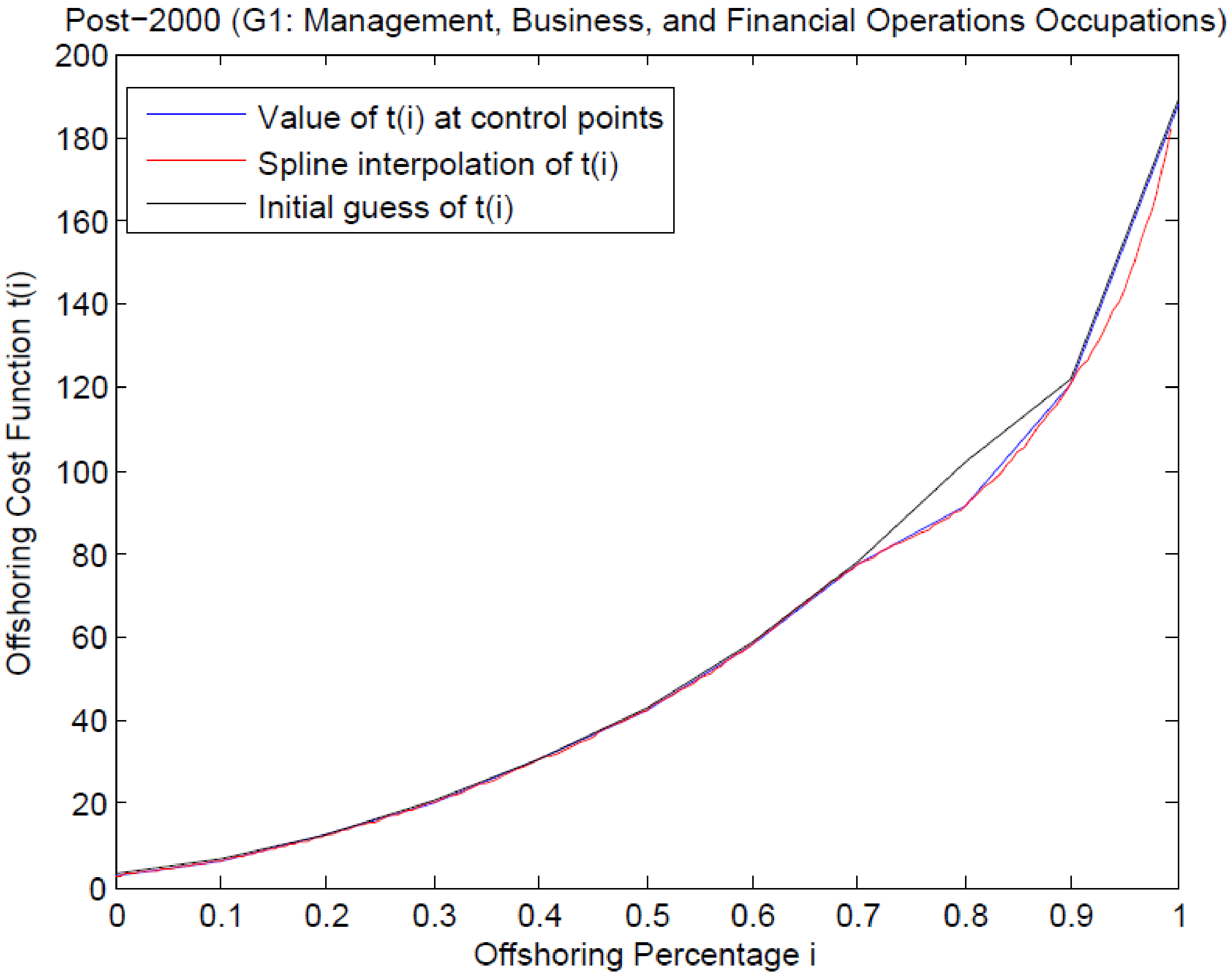


Figure 3.2.2 (cont'd)

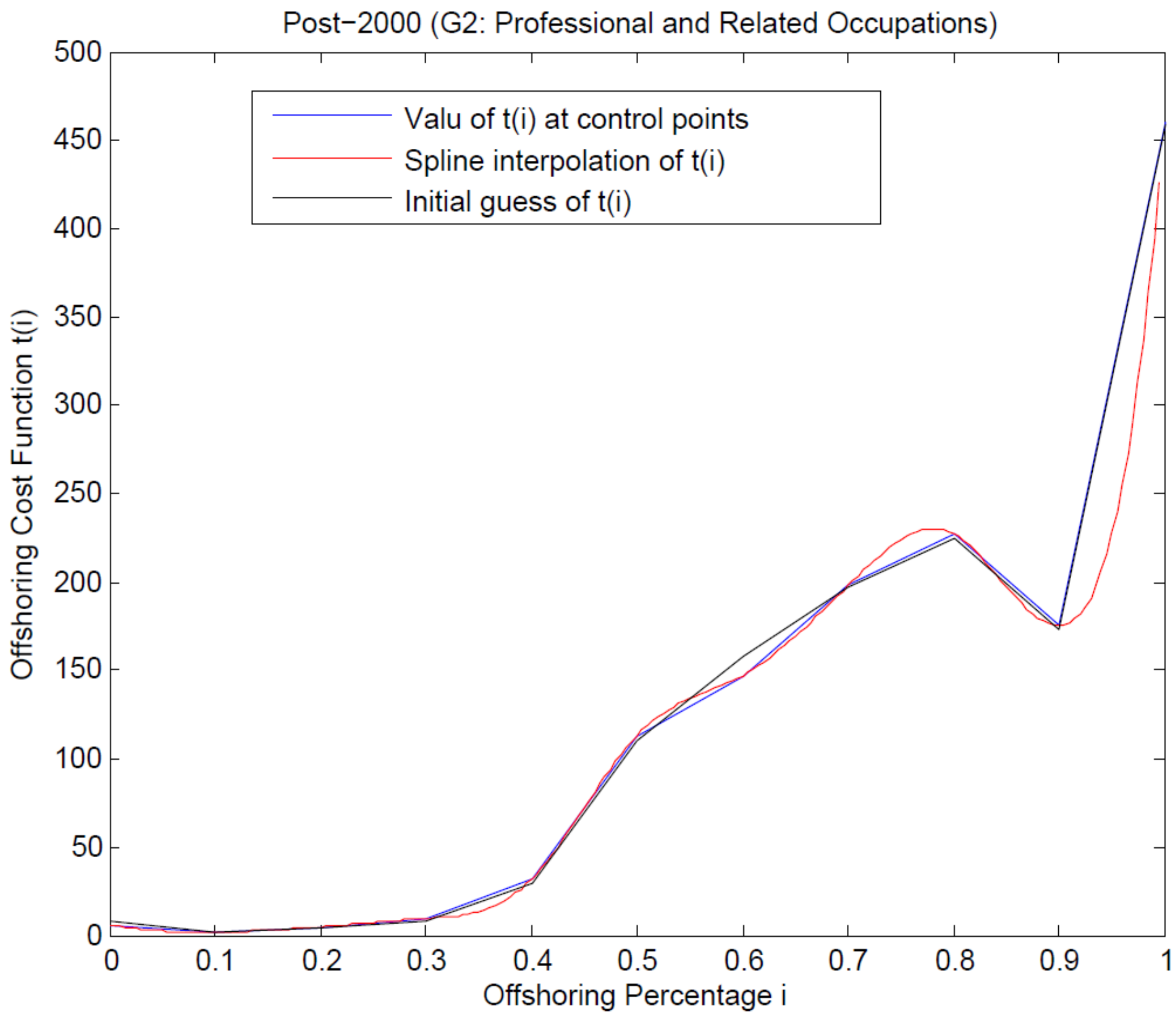


Figure 3.2.2 (cont'd)

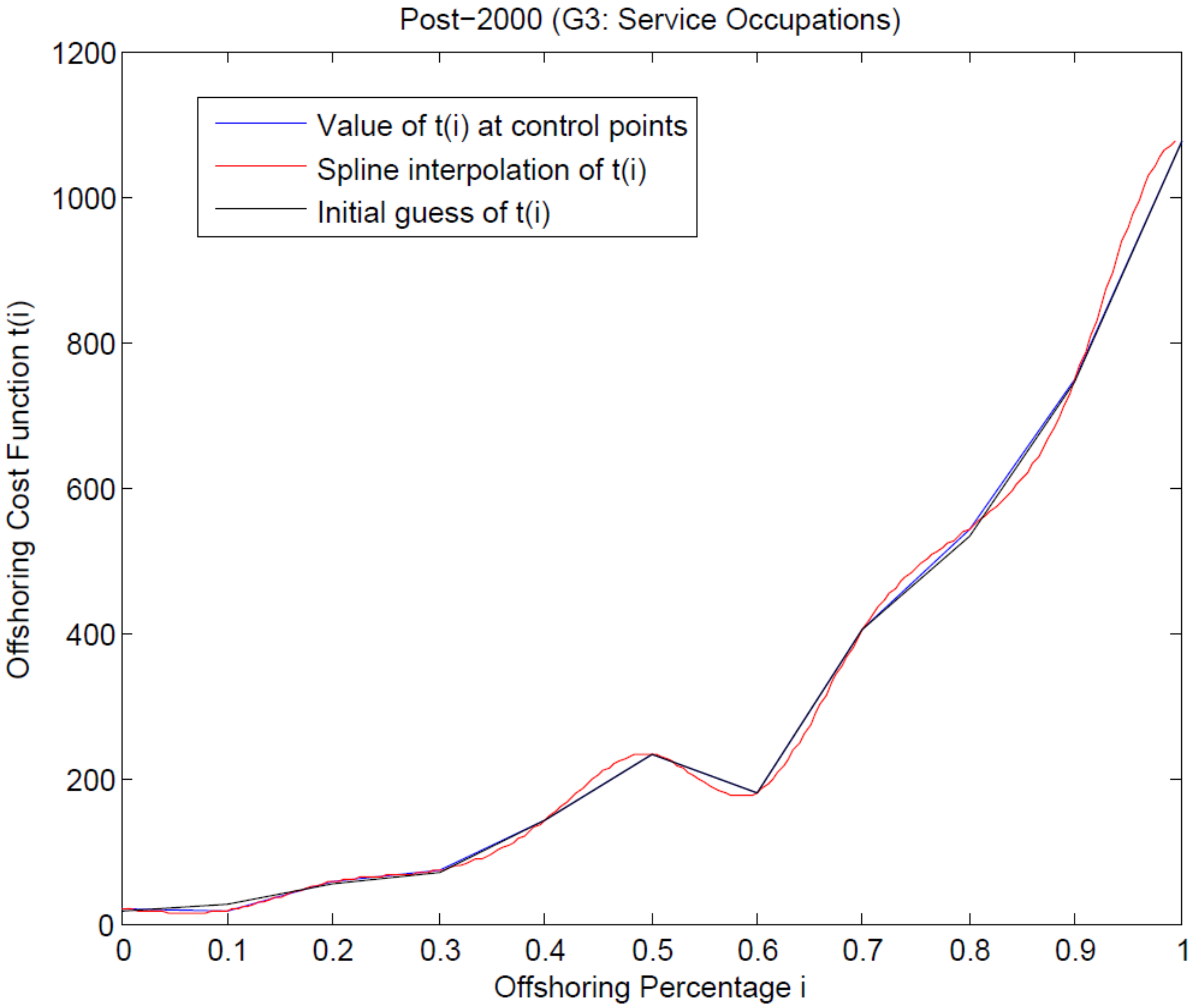


Figure 3.2.2 (cont'd)

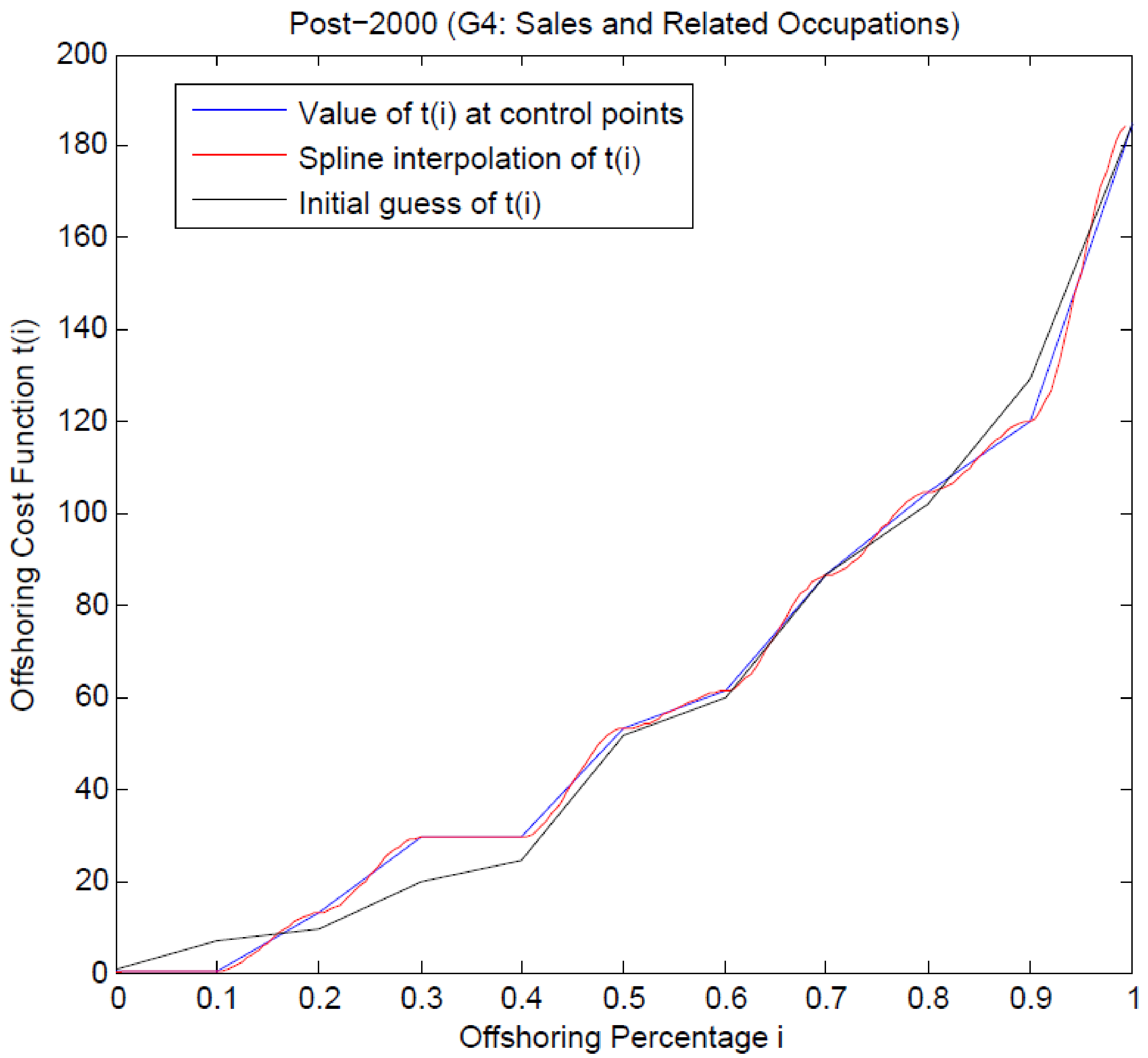


Figure 3.2.2 (cont’d)

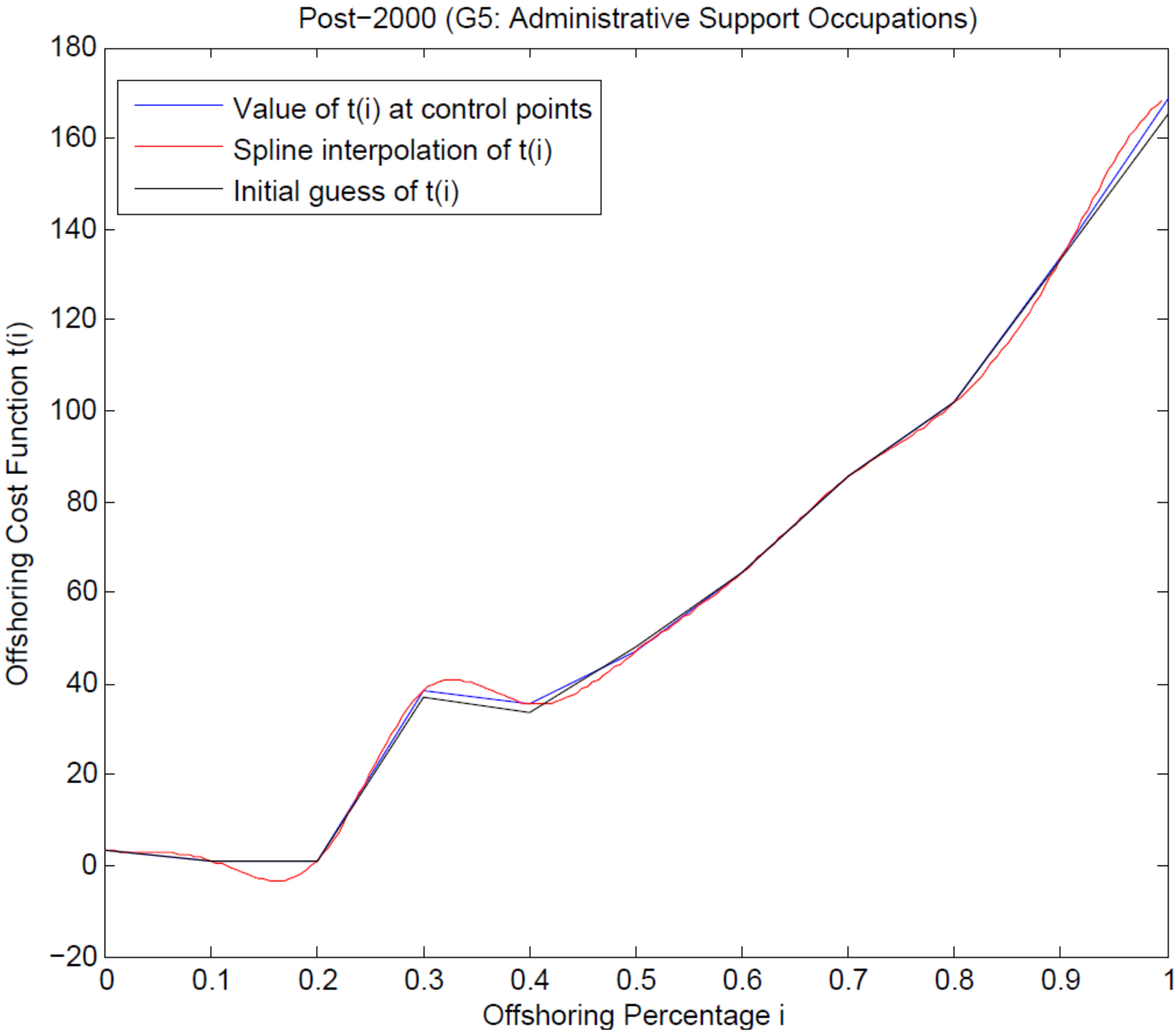


Figure 3.2.2 (cont'd)

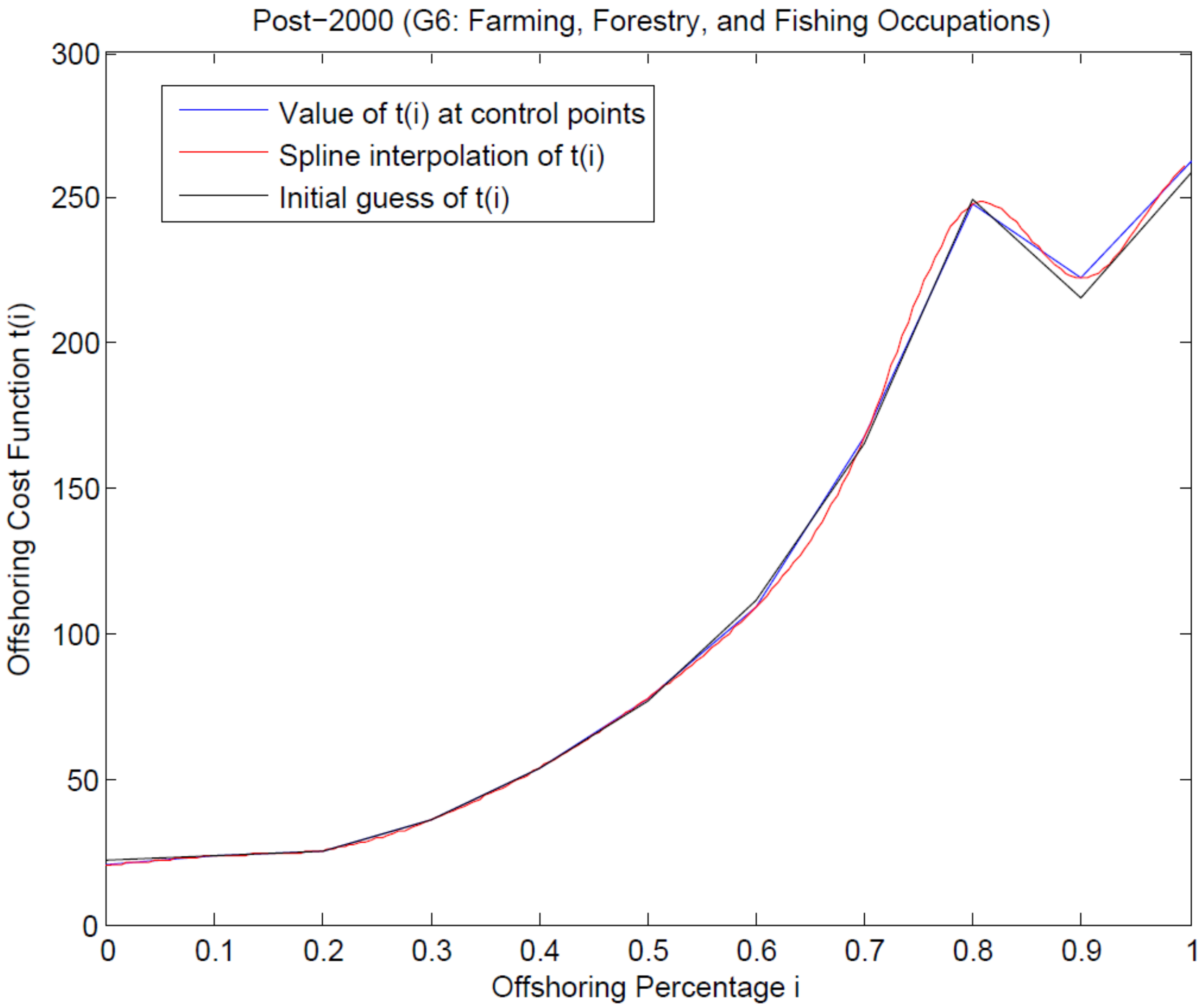


Figure 3.2.2 (cont'd)

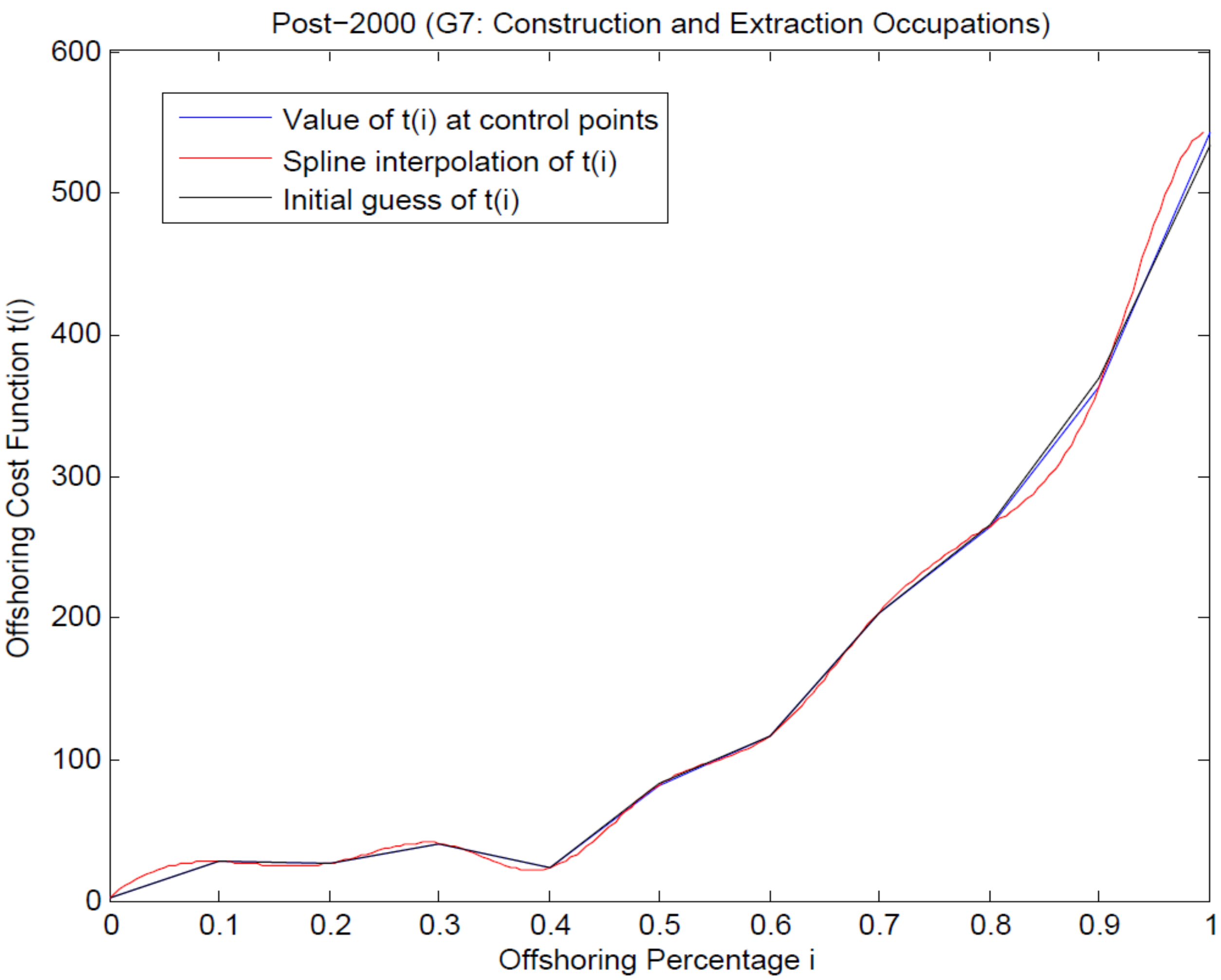


Figure 3.2.2 (cont'd)

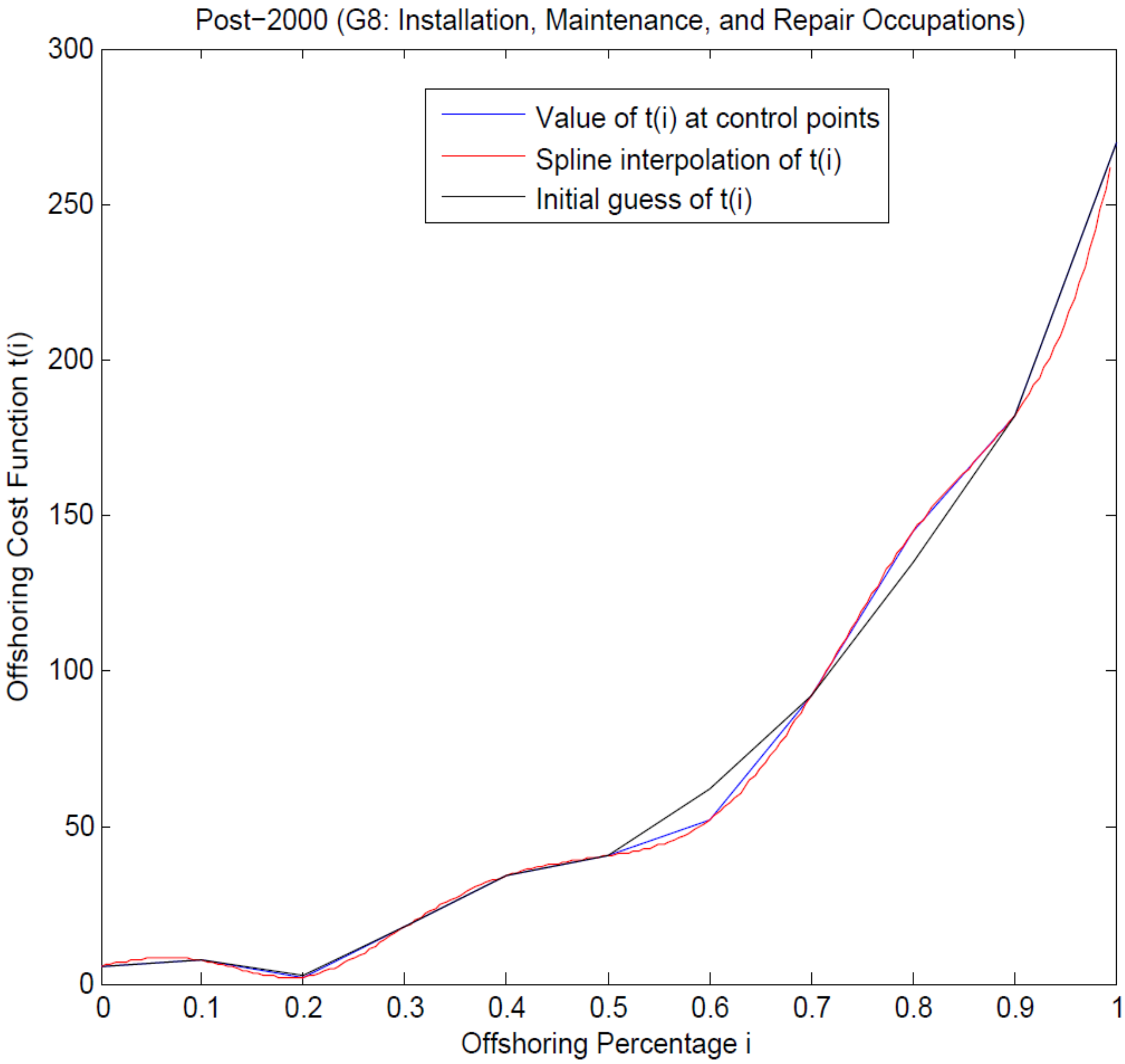


Figure 3.2.2 (cont'd)

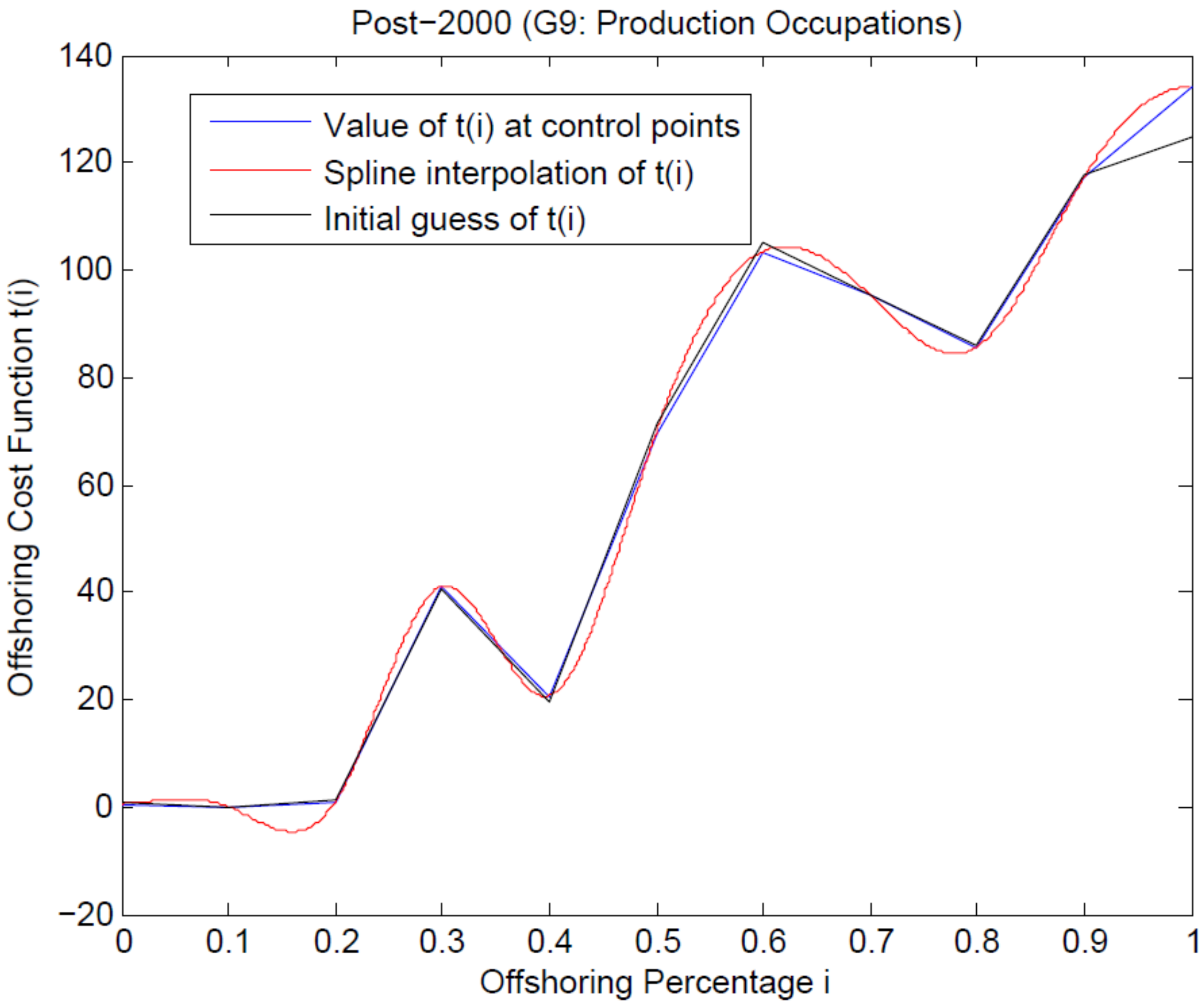


Figure 3.2.2 (cont'd)

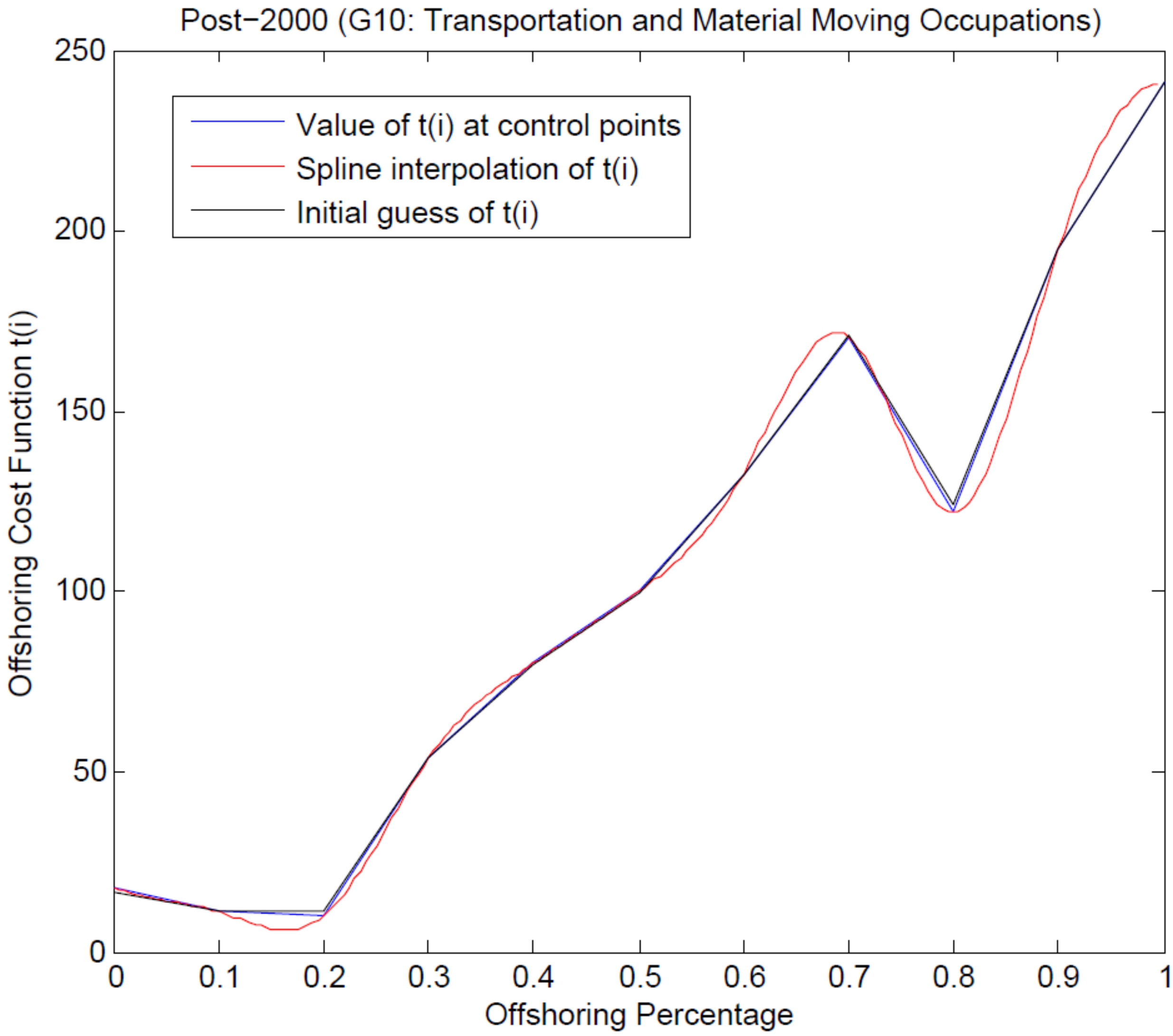


Figure 3.3.1: NLS Method Cubic Offshoring Cost Function for G1 (Management, Business and Financial Operations Occupations)

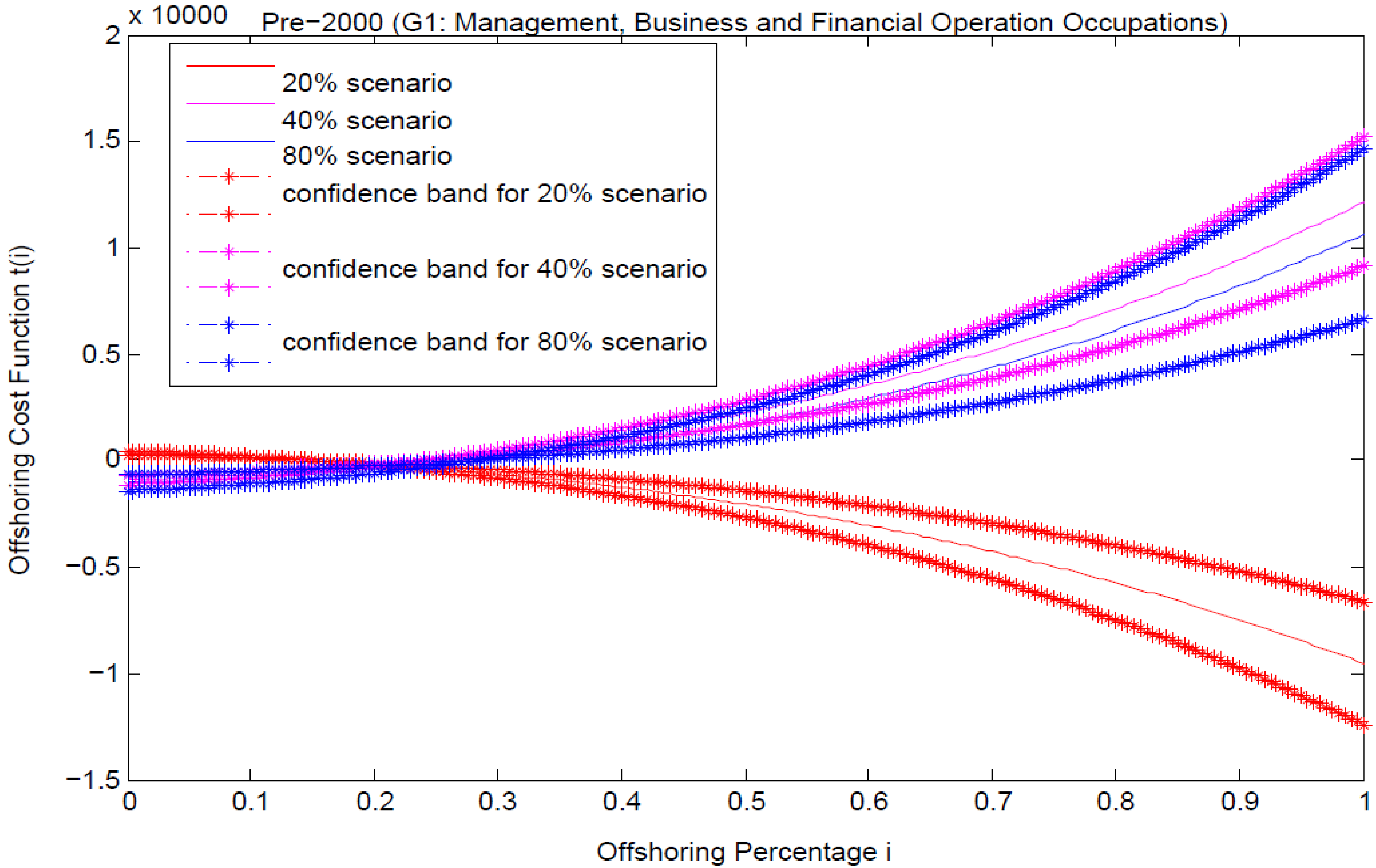
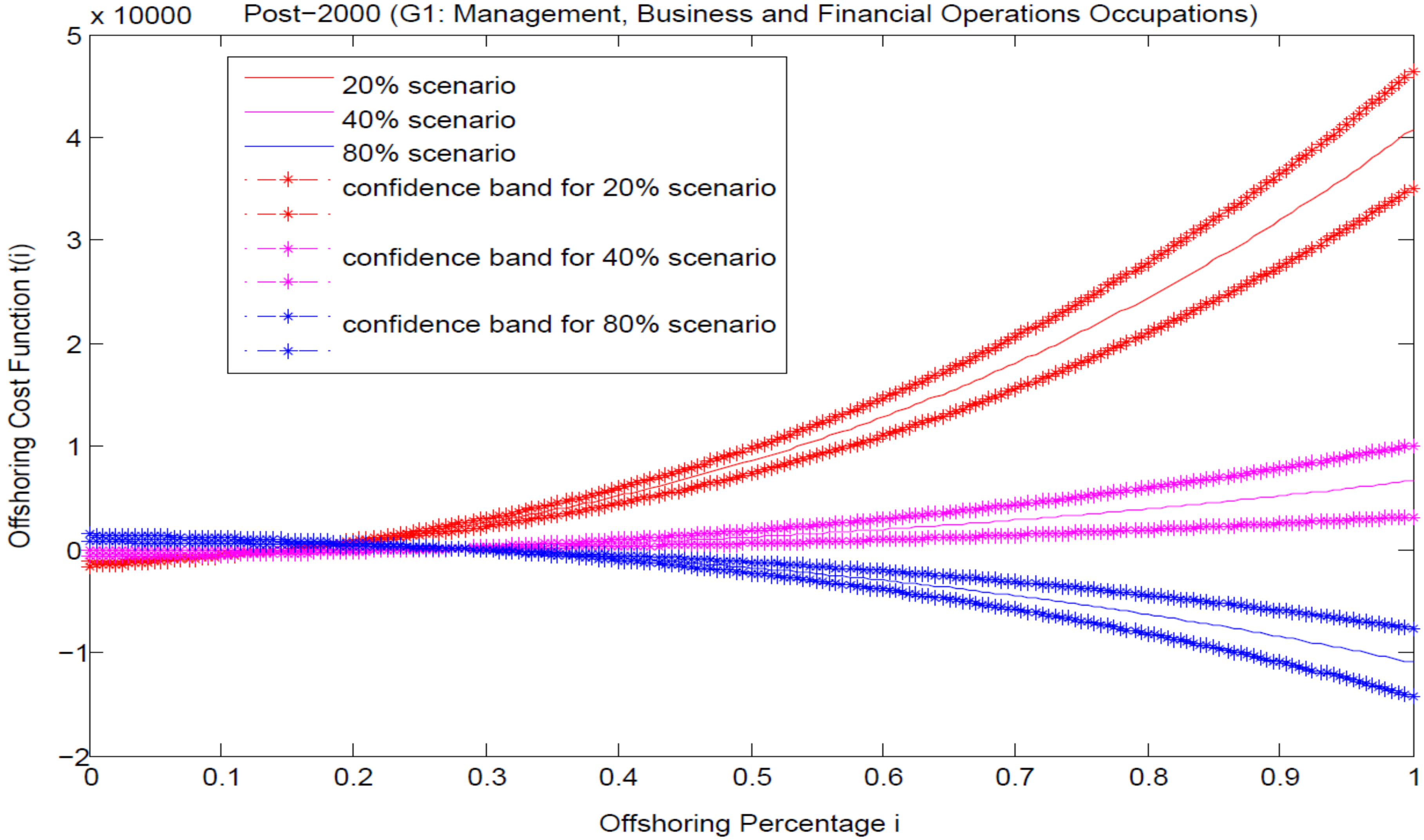


Figure 3.3.1 (cont'd)



Notes: 95 percent confidence band is calculated with 50 times bootstrapping.

Figure 3.3.2: NLS Method Cubic Offshoring Cost Function for G2 (Professional and Related Occupations)

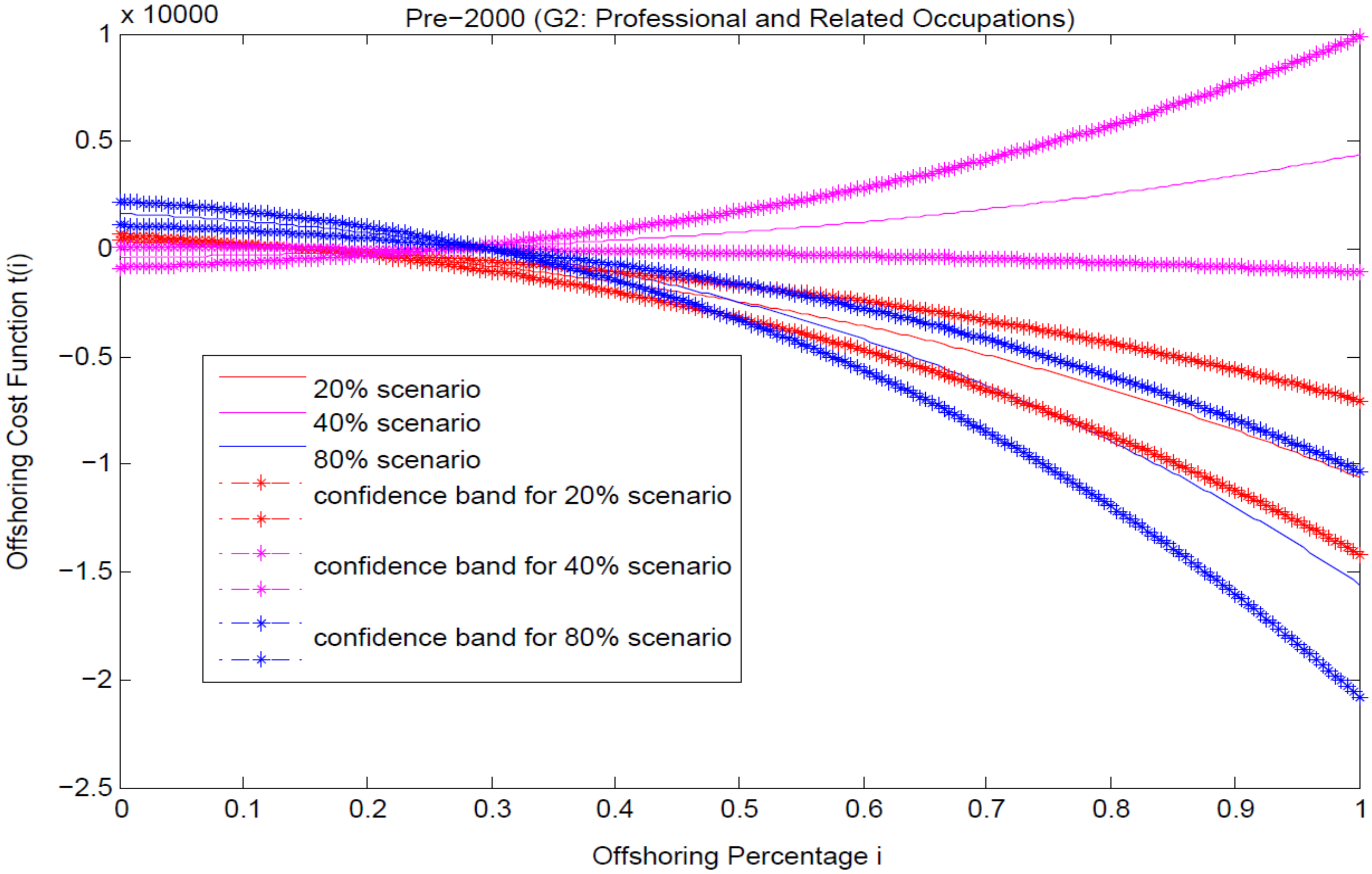
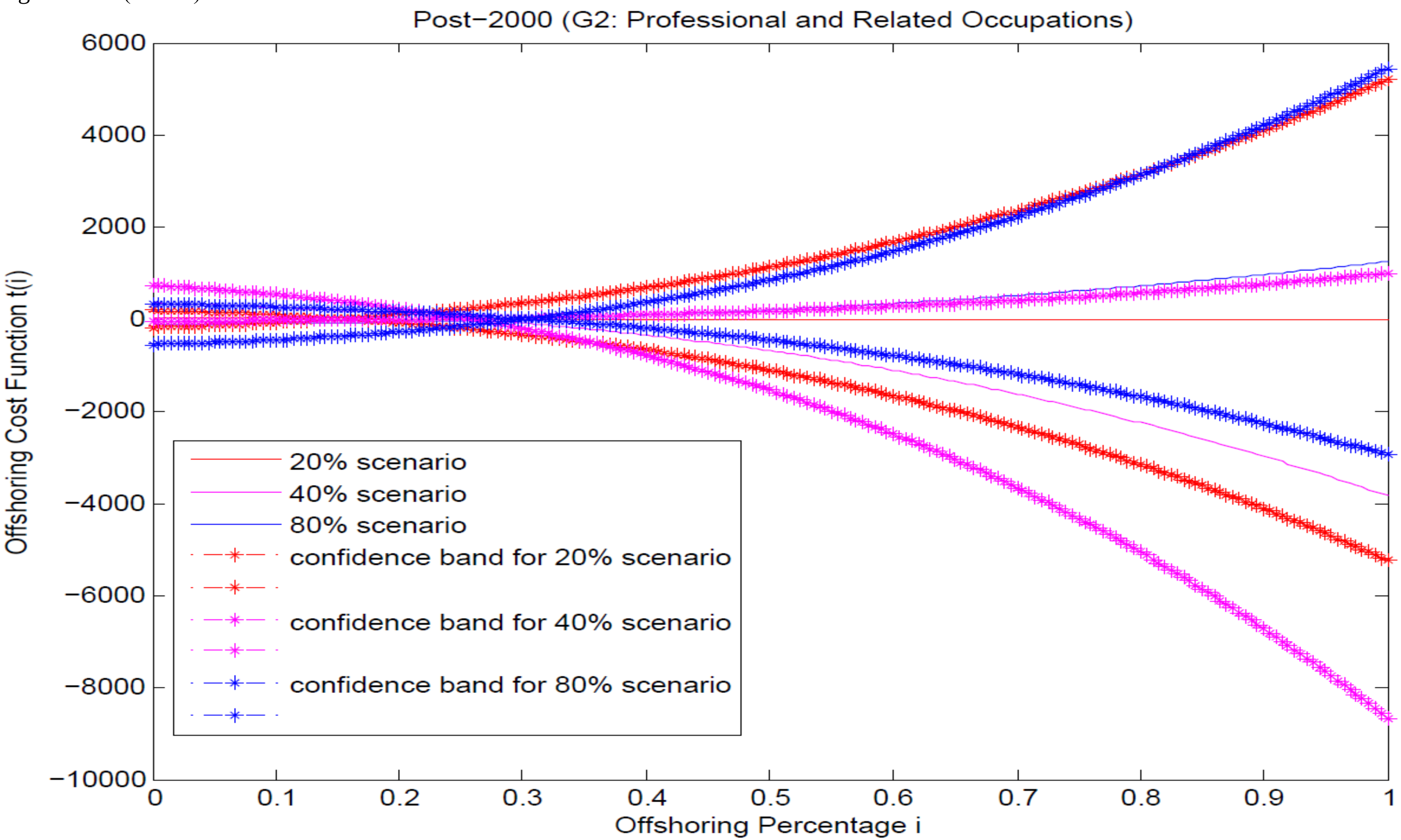


Figure 3.3.2 (cont'd)



Notes: 95 percent confidence band is calculated with 50 times bootstrapping.

Figure 3.3.3: NLS Method Cubic Offshoring Cost Function for G4 (Sales and Related Occupations)

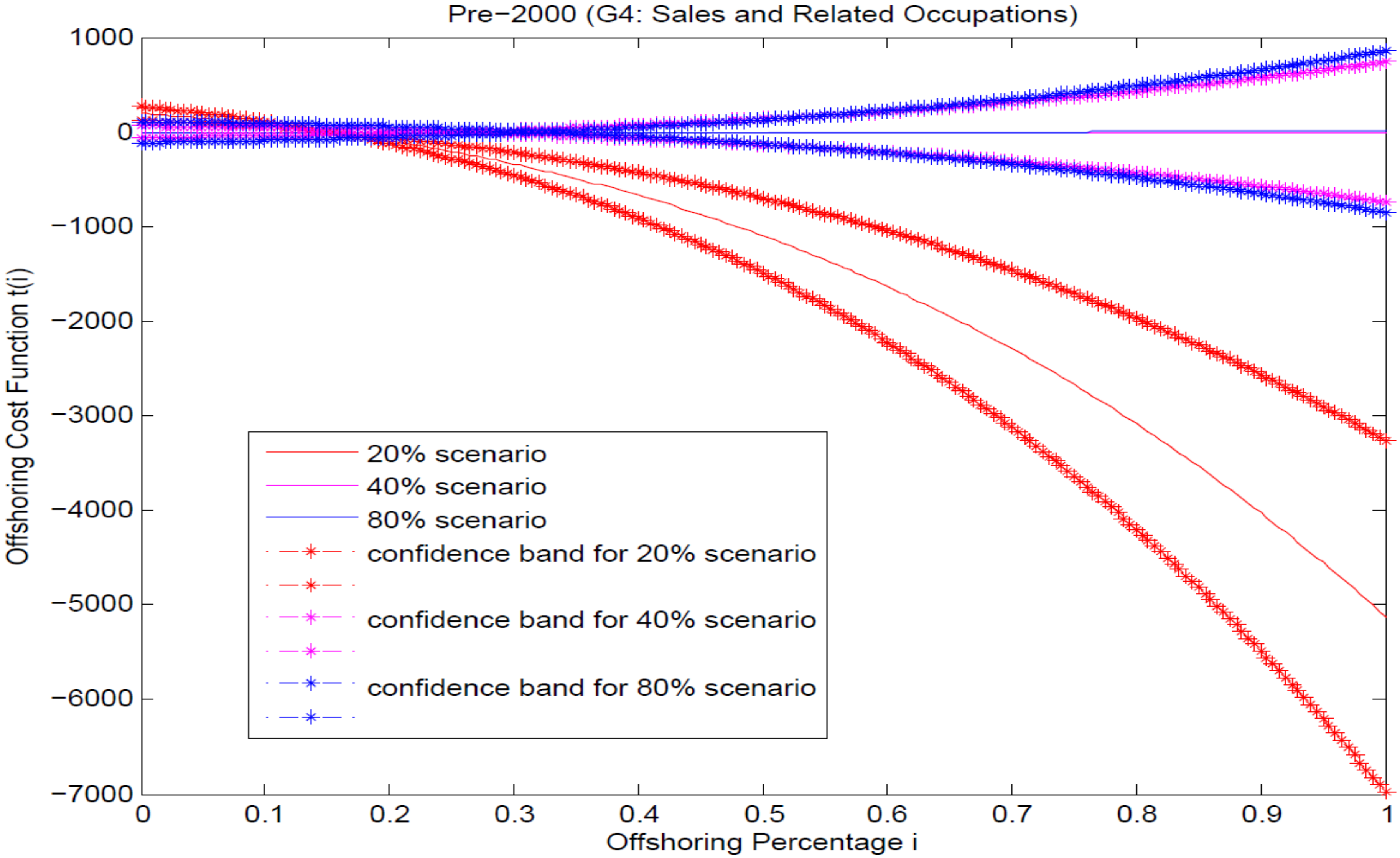
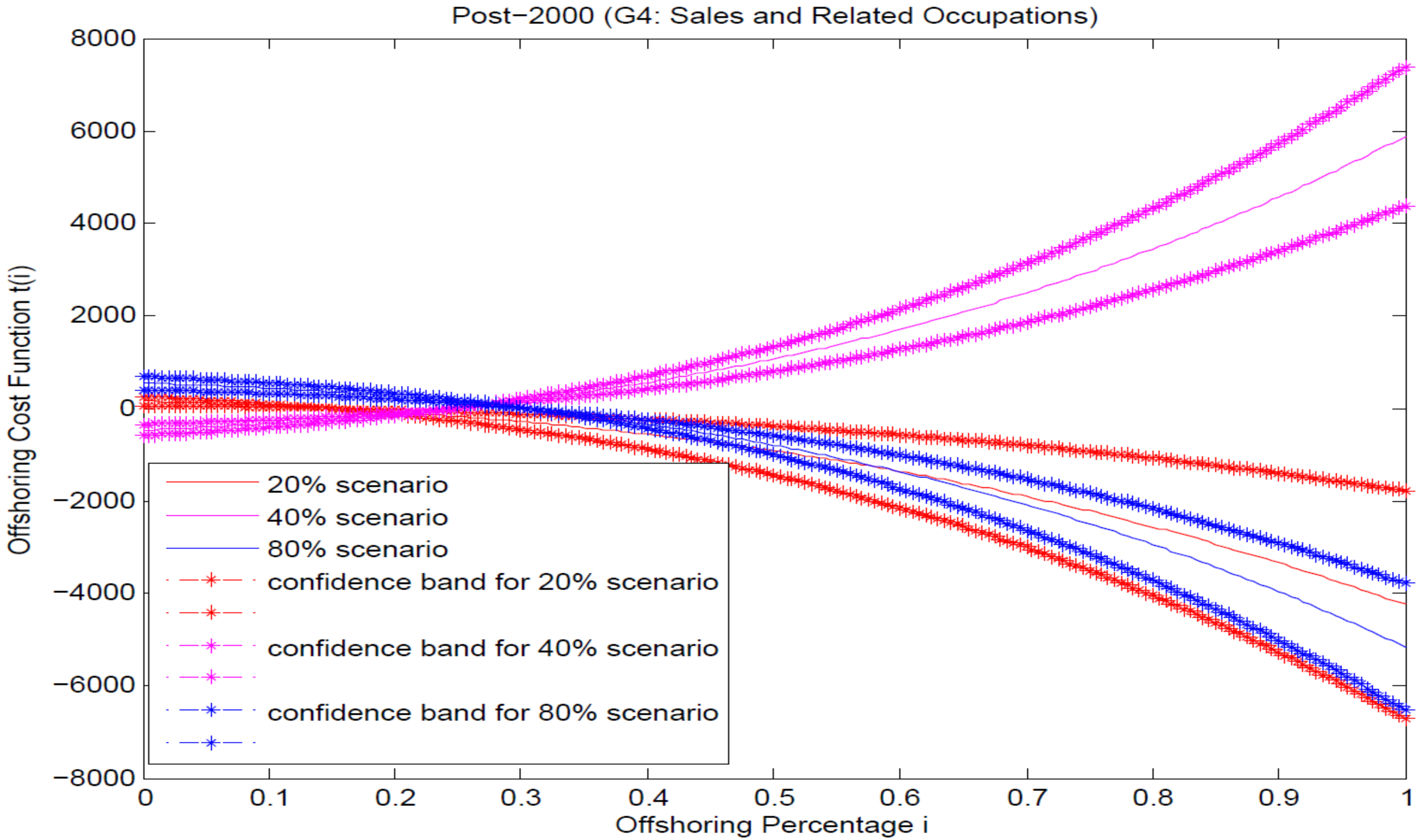


Figure 3.3.3 (cont'd)



Notes: 95 percent confidence band is calculated with 50 times bootstrapping.

Figure 3.3.4: NLS Method Cubic Offshoring Cost Function for G5 (Office and Administrative Support Occupations)

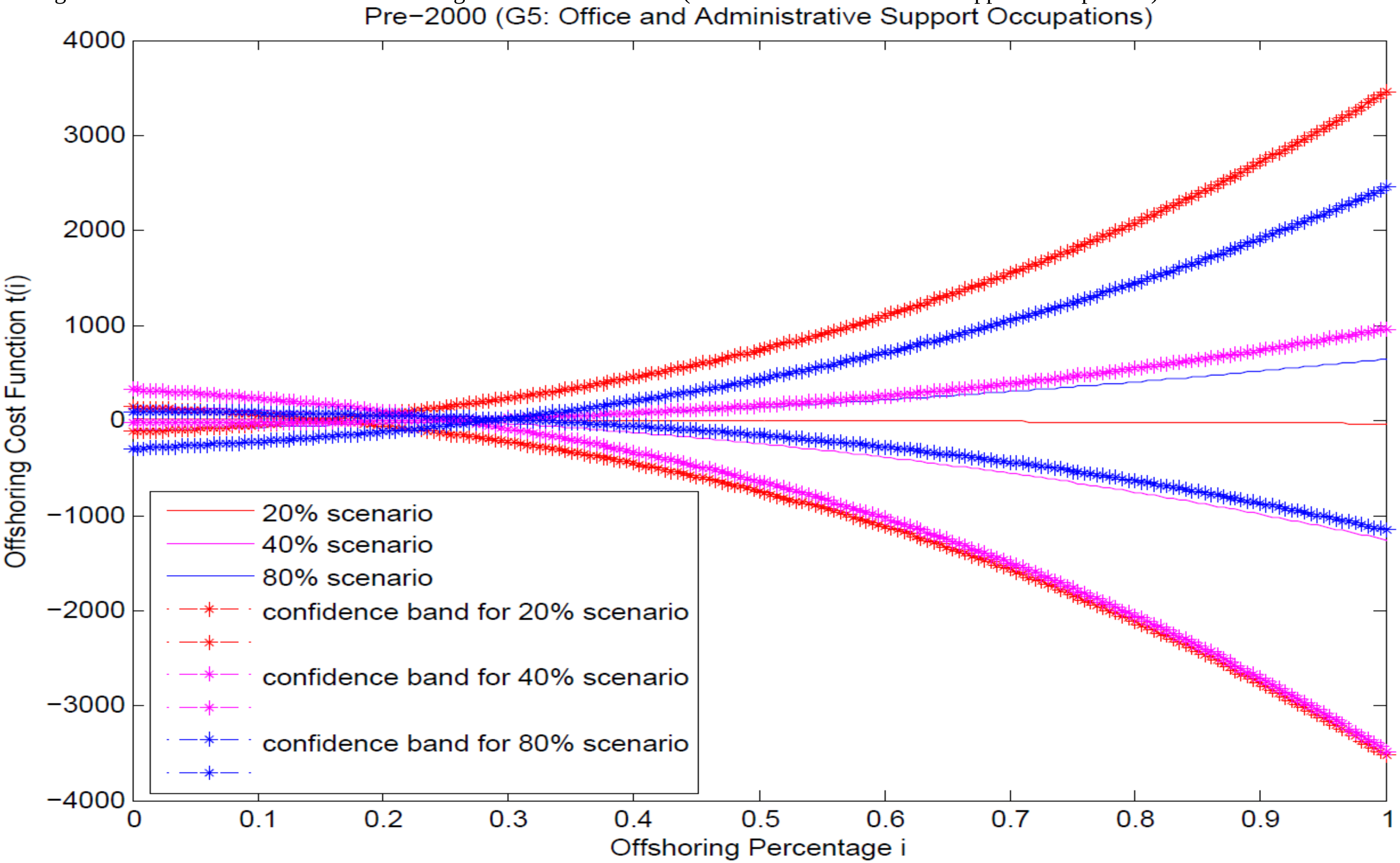
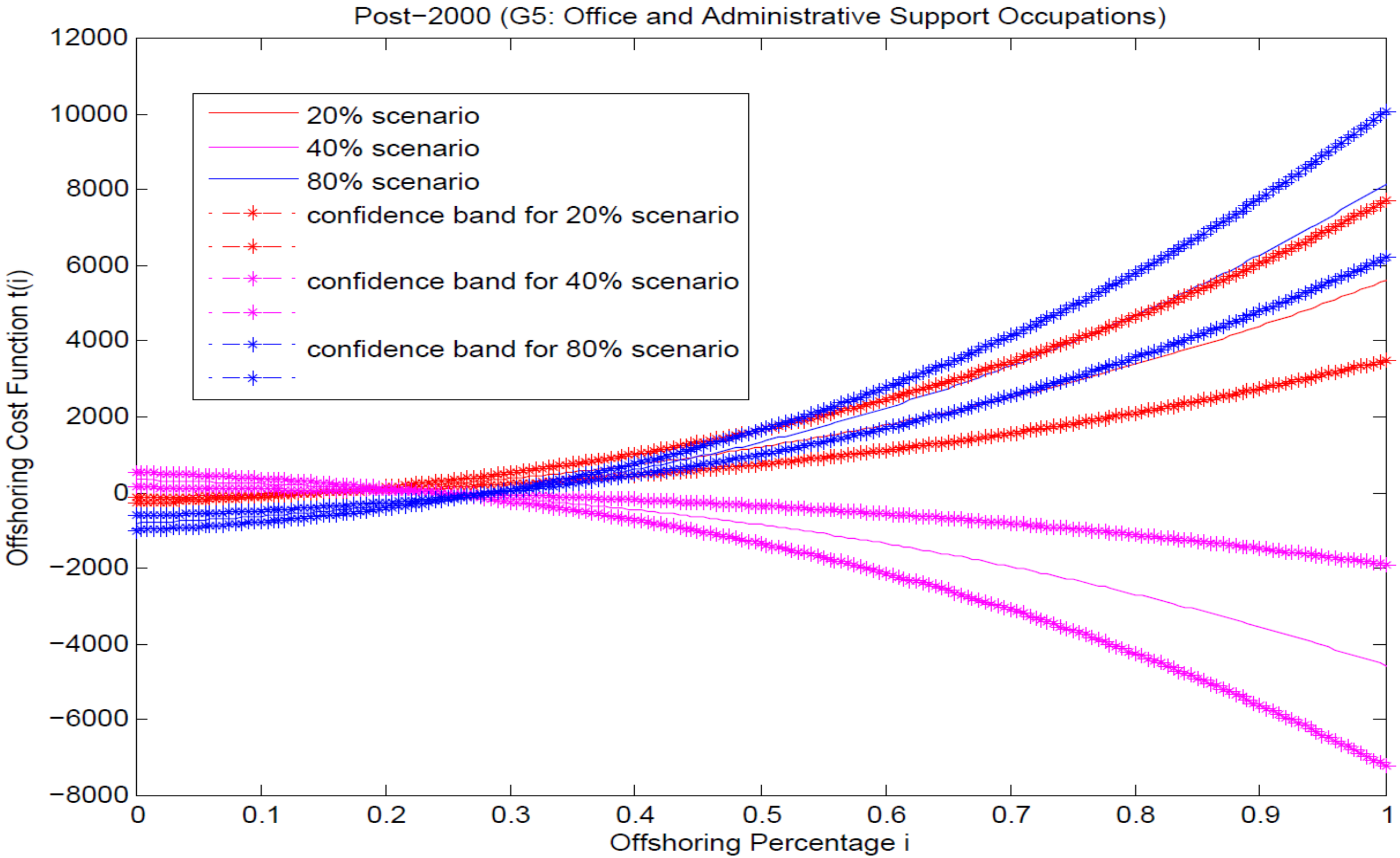


Figure 3.3.4 (cont'd)



Notes: 95 percent confidence band is calculated with 50 times bootstrapping.

Figure 3.3.5: NLS Method Cubic Offshoring Cost Function for G9 (Production Occupations)

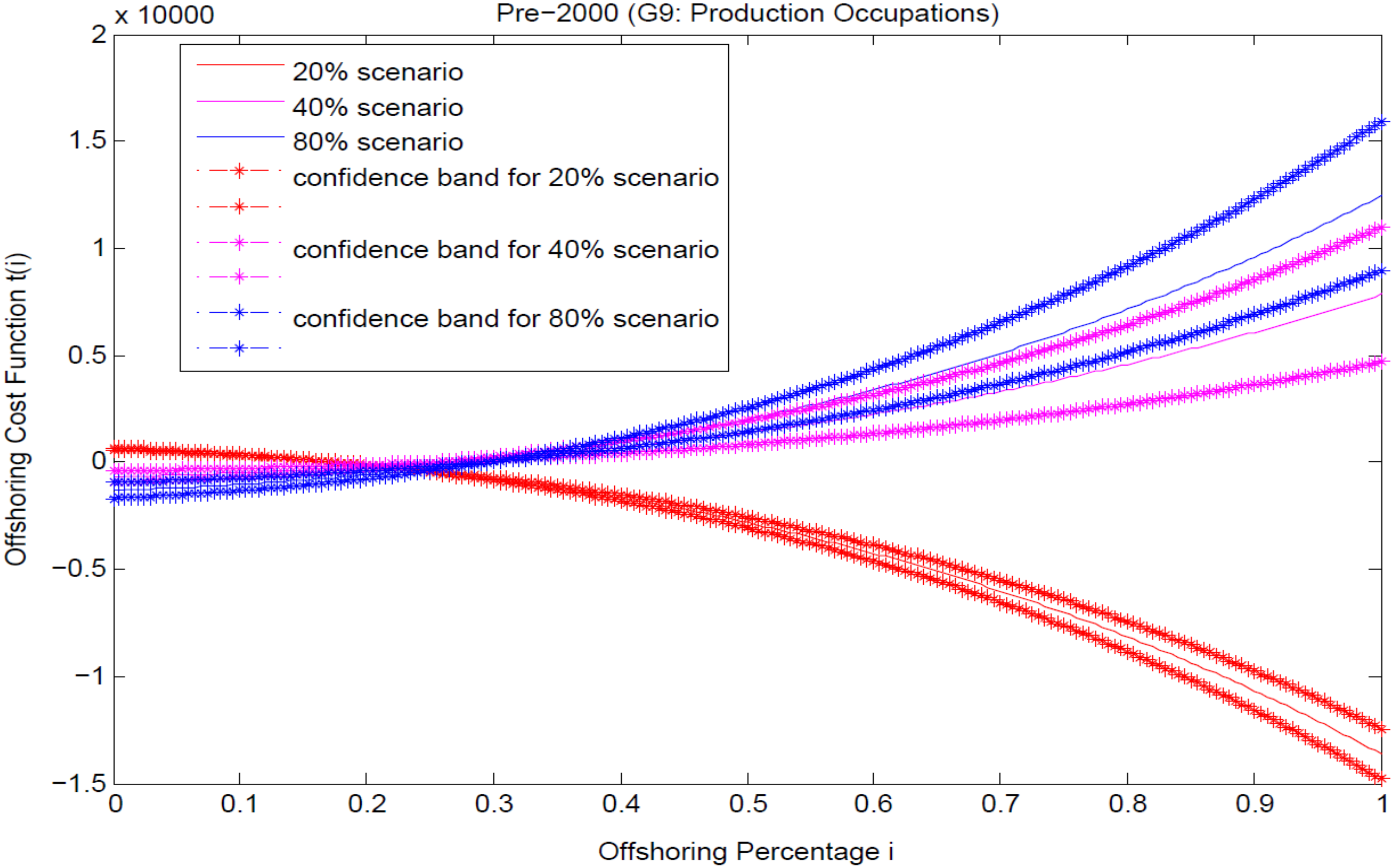
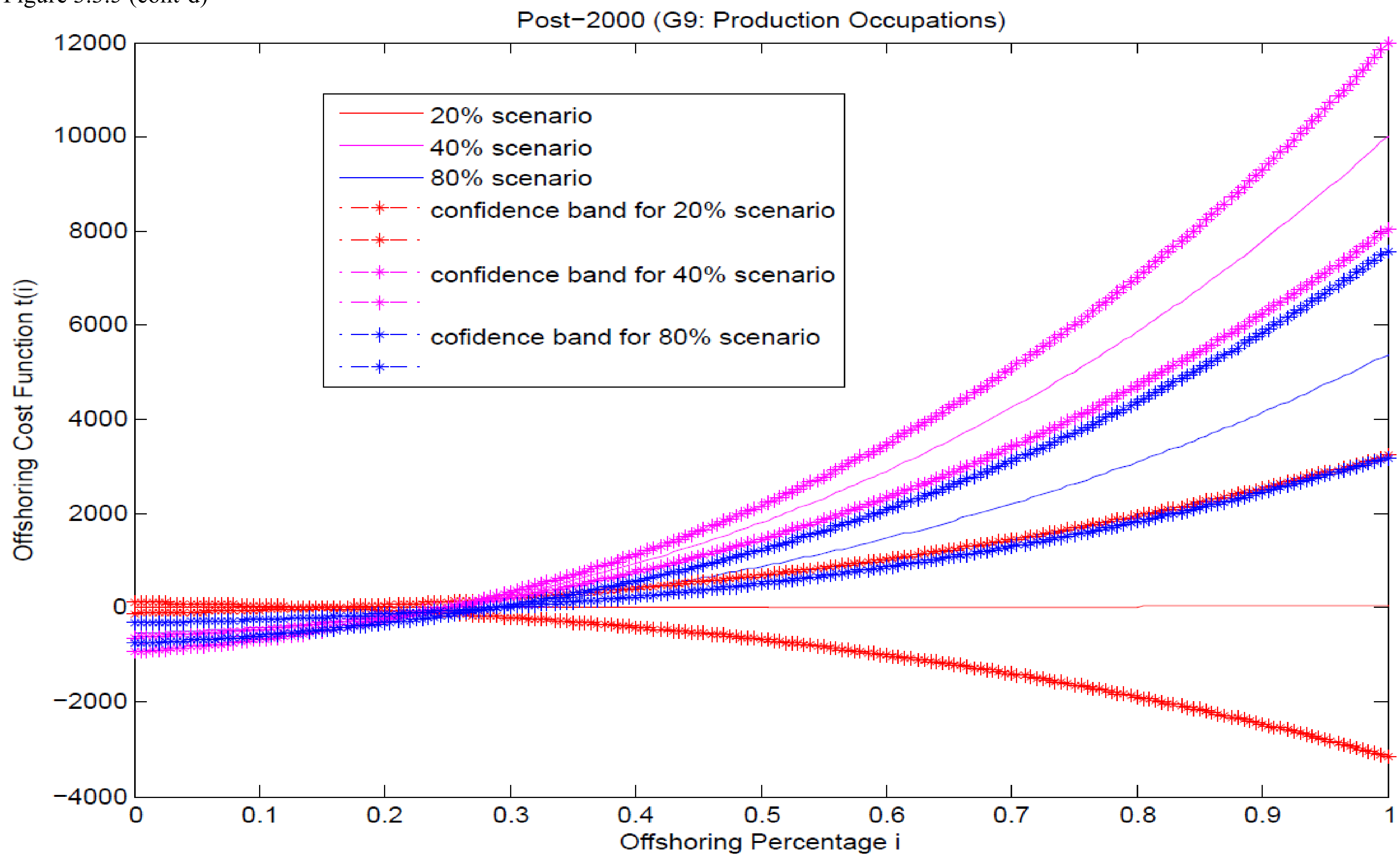
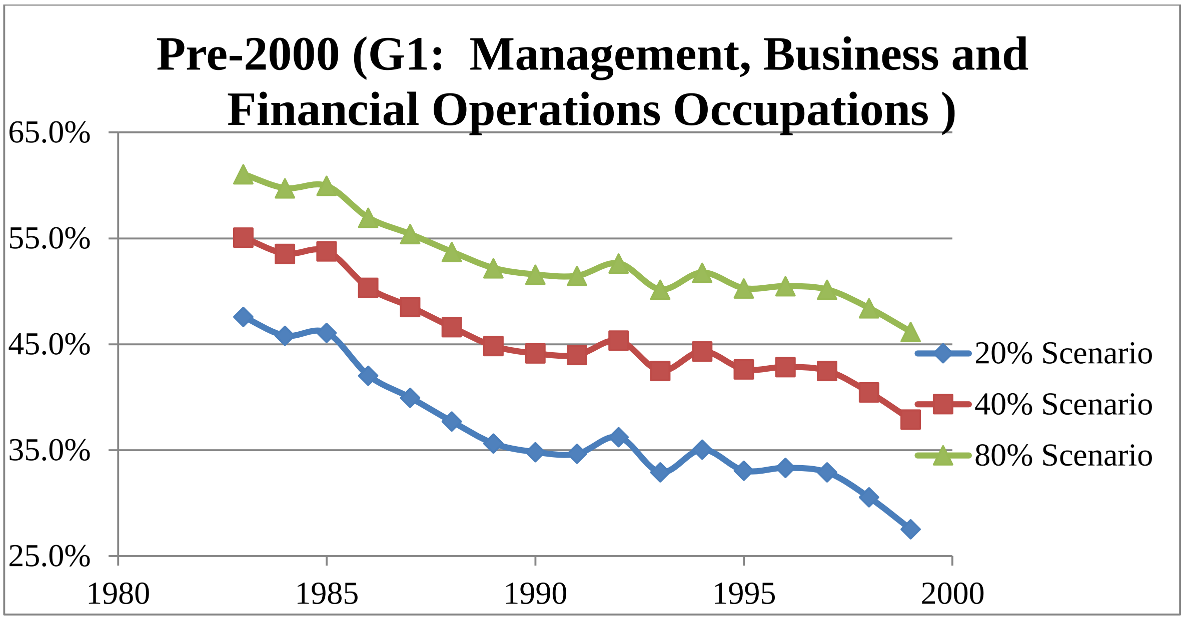


Figure 3.3.5 (cont'd)



Notes: 95 percent confidence band is calculated with 50 times bootstrapping.

Figure 3.4.1: Change of Offshoring Percentage for G1 (Management, Business and Financial Operations Occupations)



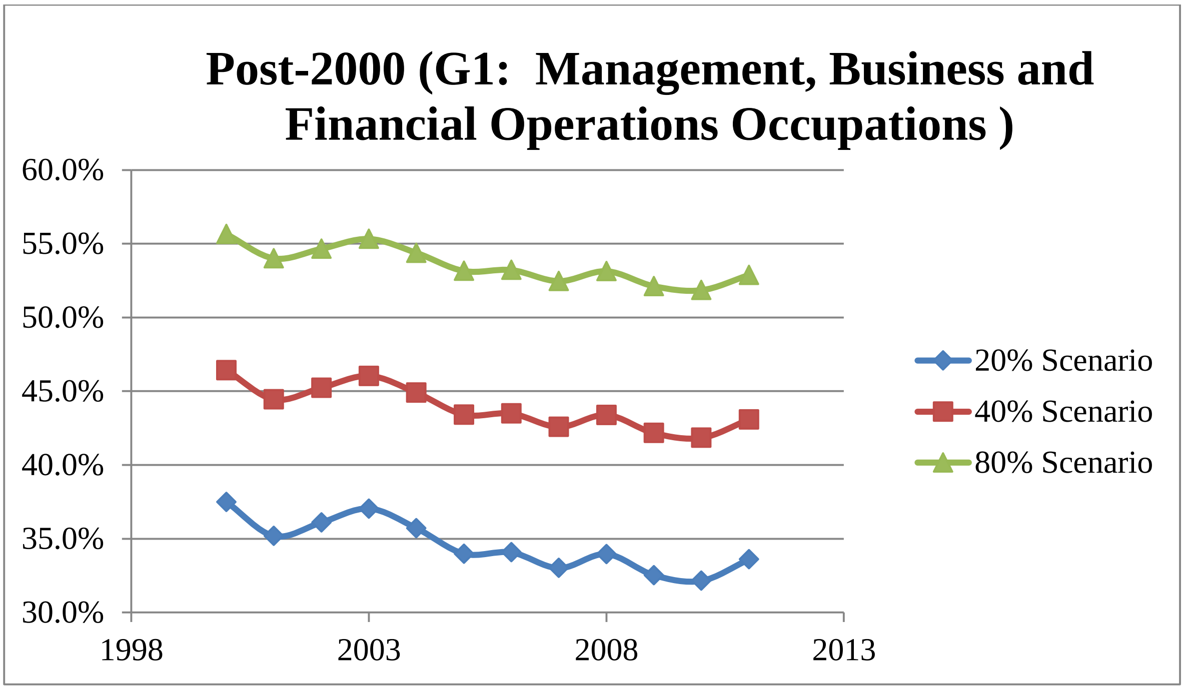
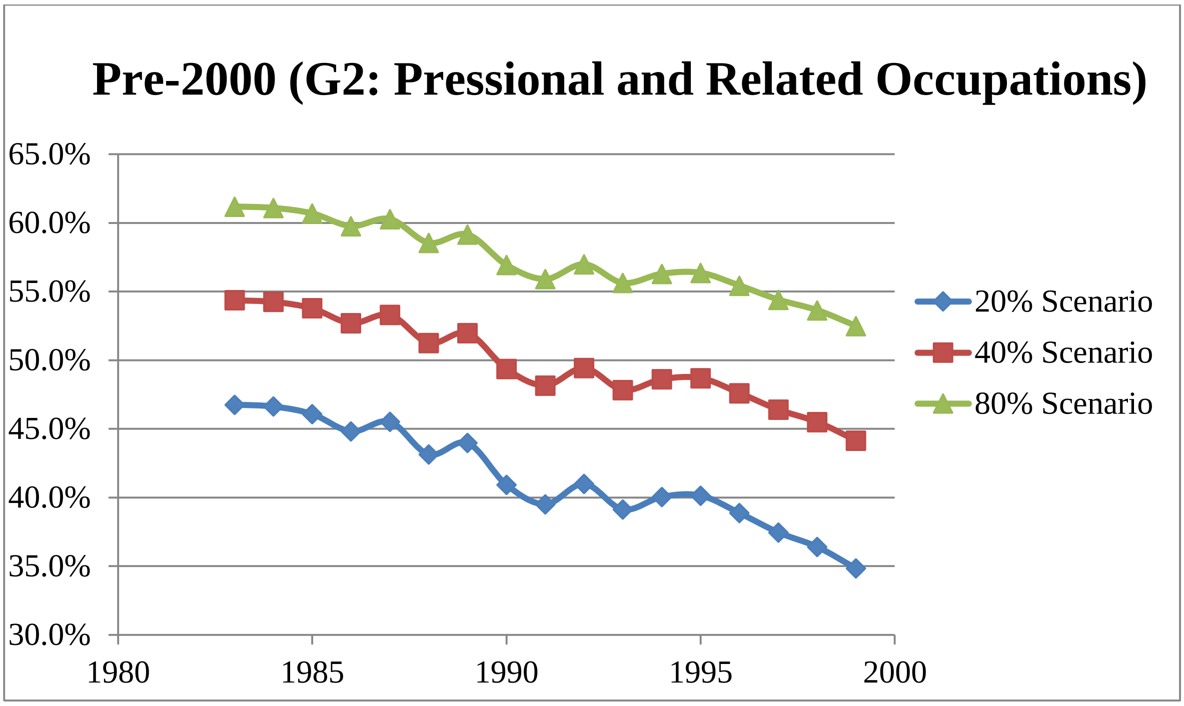


Figure 3.4.2: Change of Offshoring Percentage for G2 (Professional and Related Occupations)



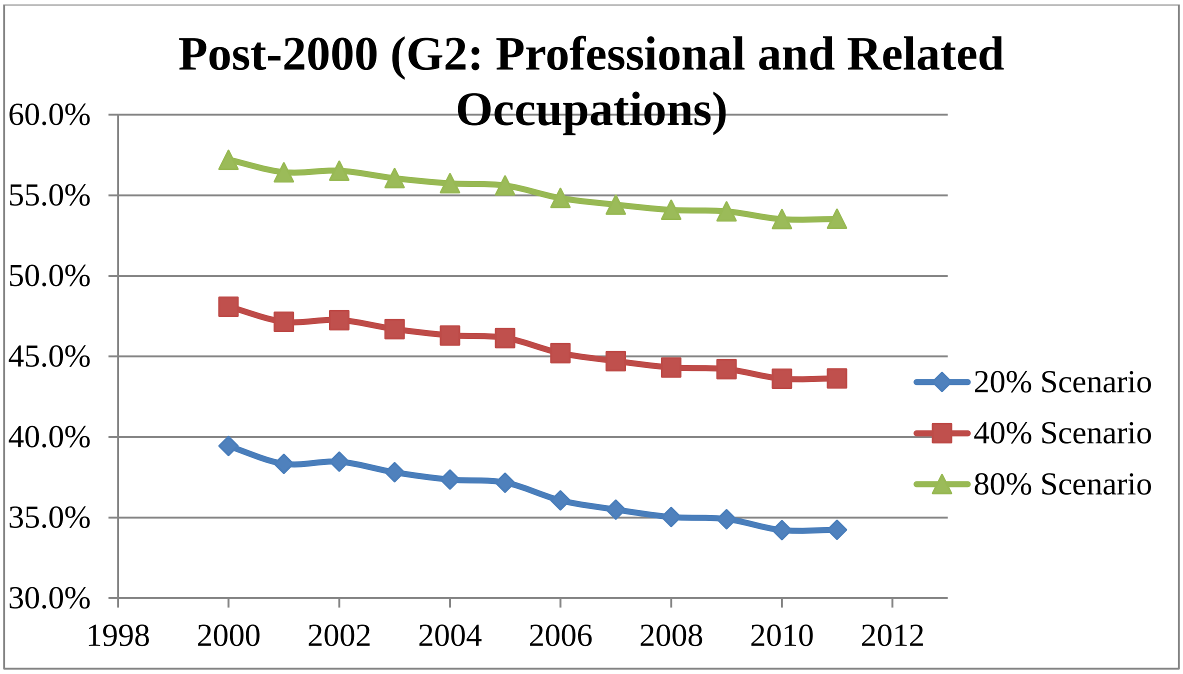
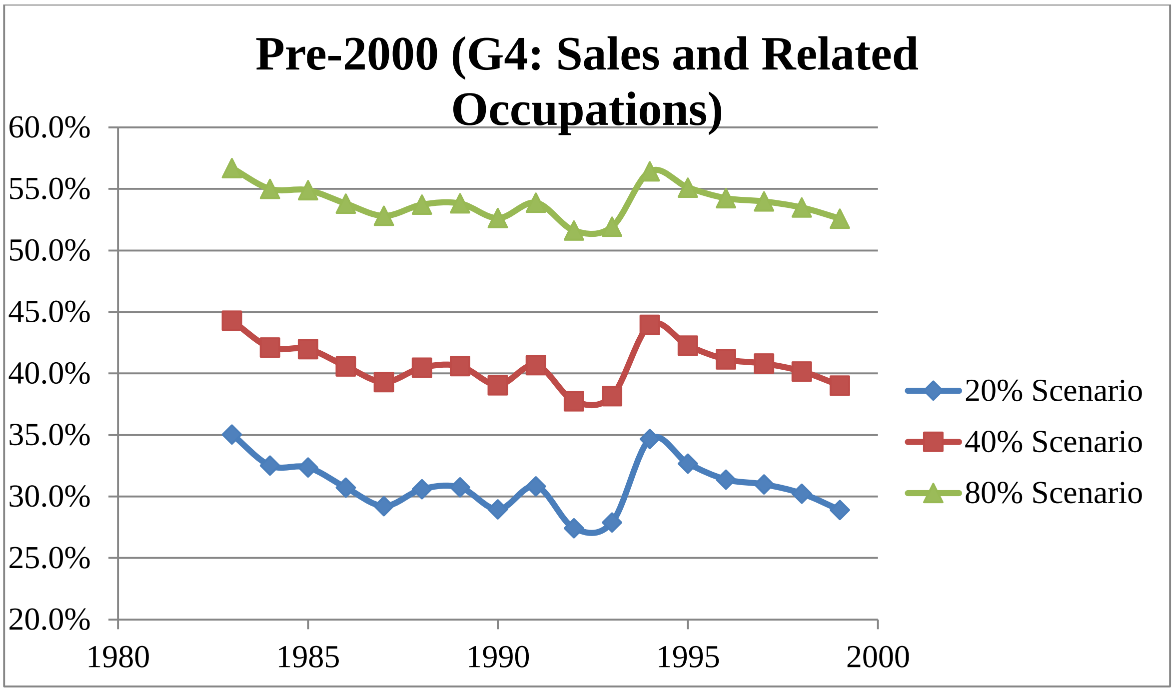


Figure 3.4.3: Change of Offshoring Percentage for G4 (Sales and Related Occupations)



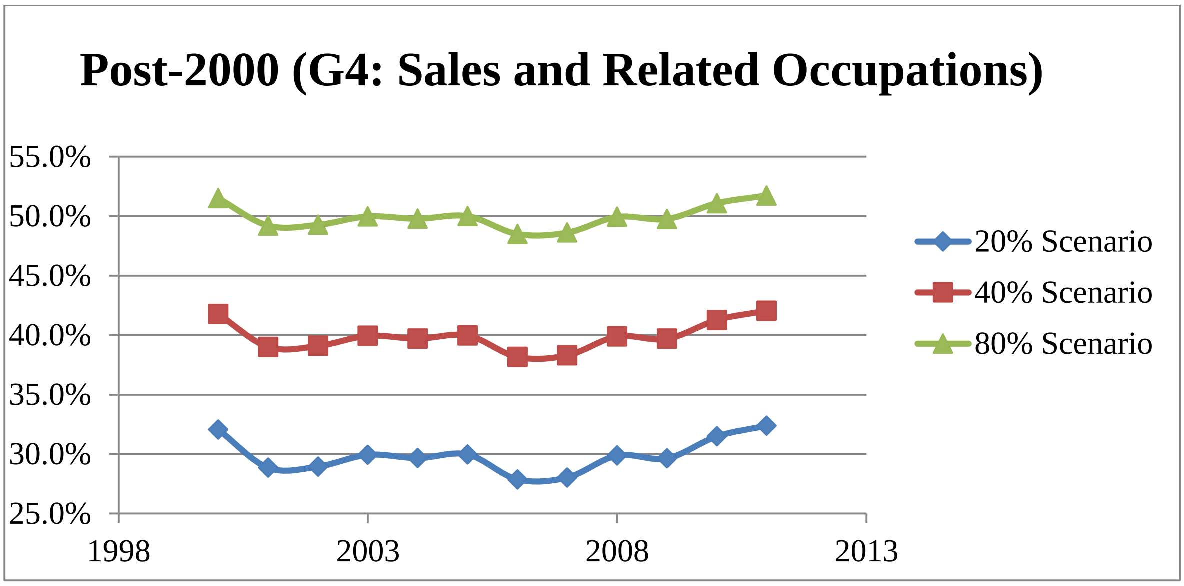
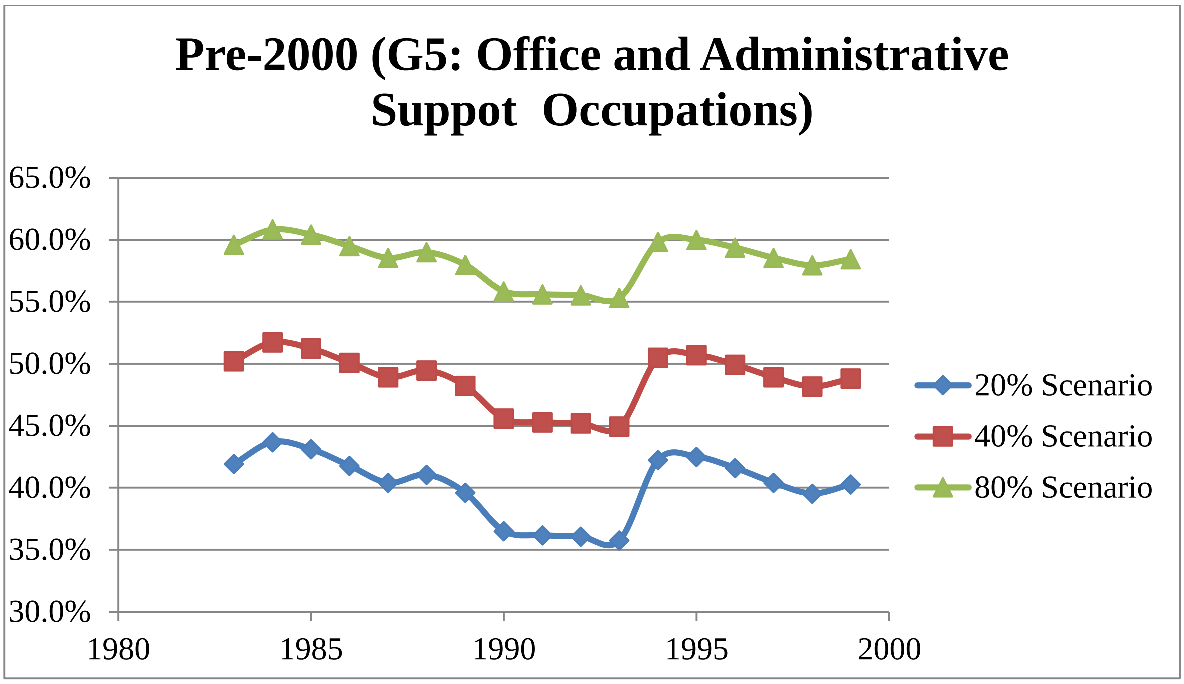


Figure 3.4.4: Change of Offshoring Percentage for G5 (Office and Administrative Support Occupations)



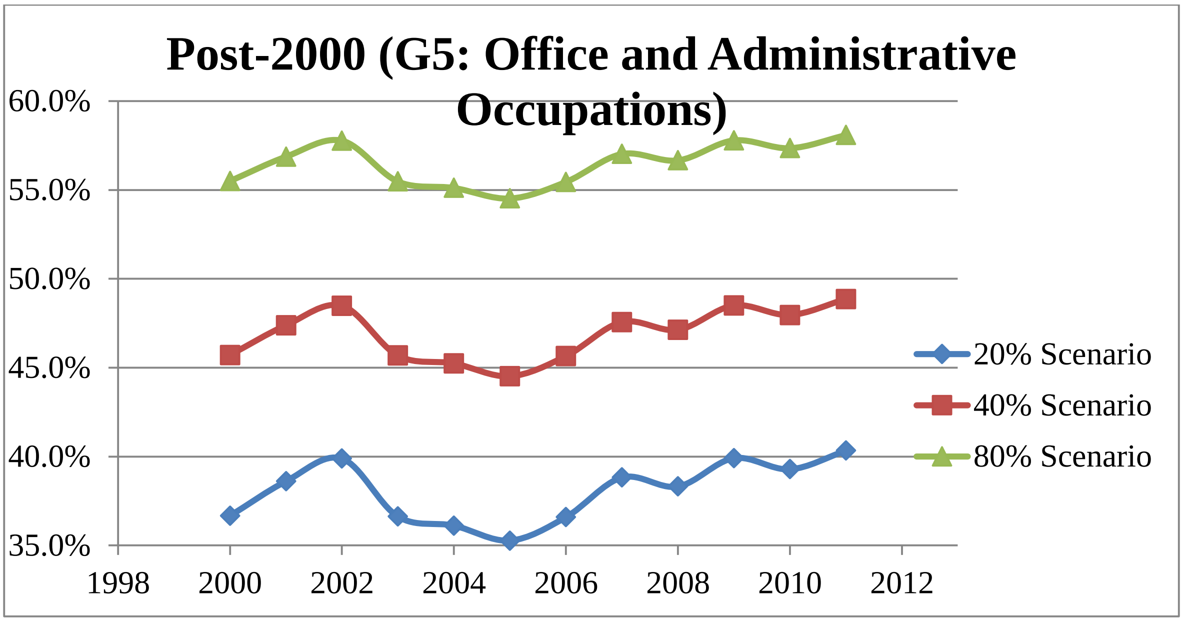
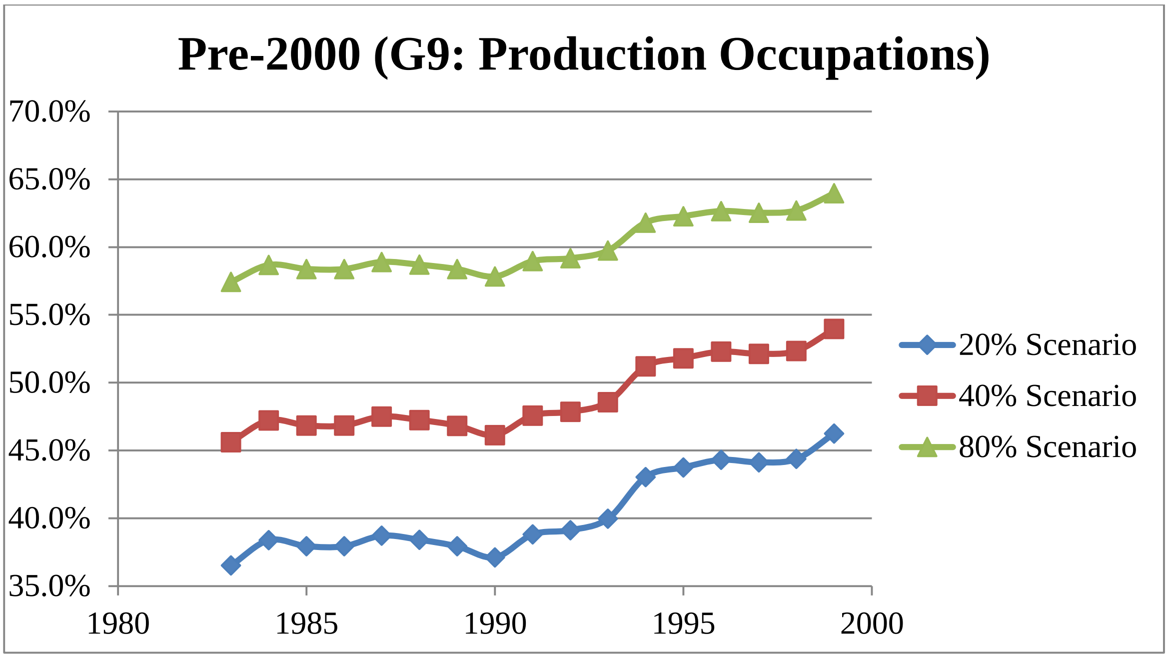


Figure 3.4.5: Change of Offshoring Percentage for G9 (Production Occupations)



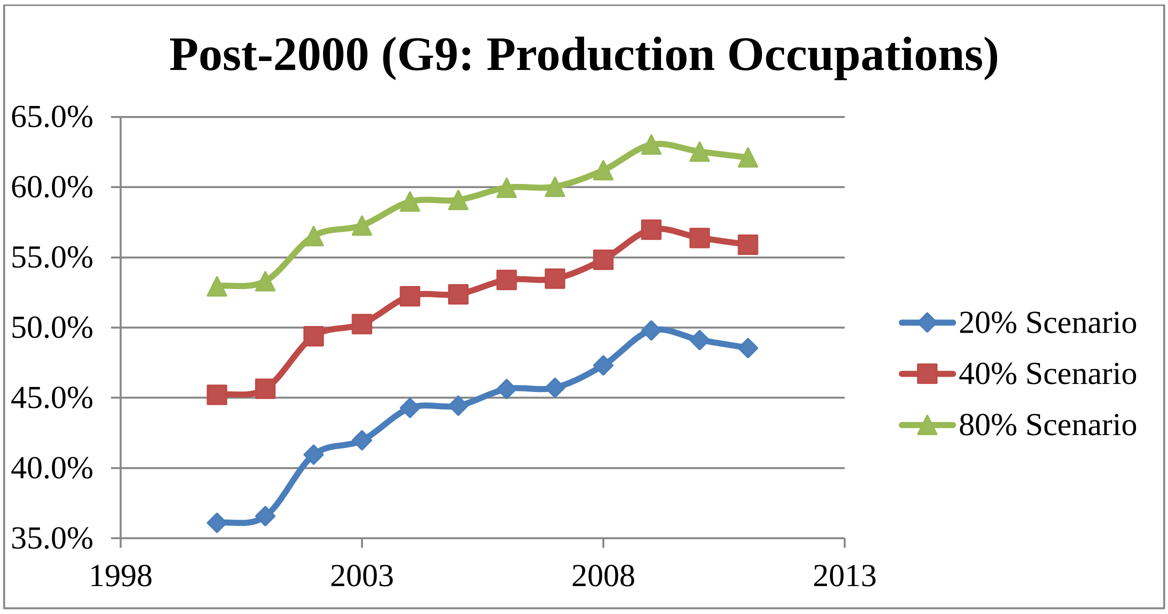
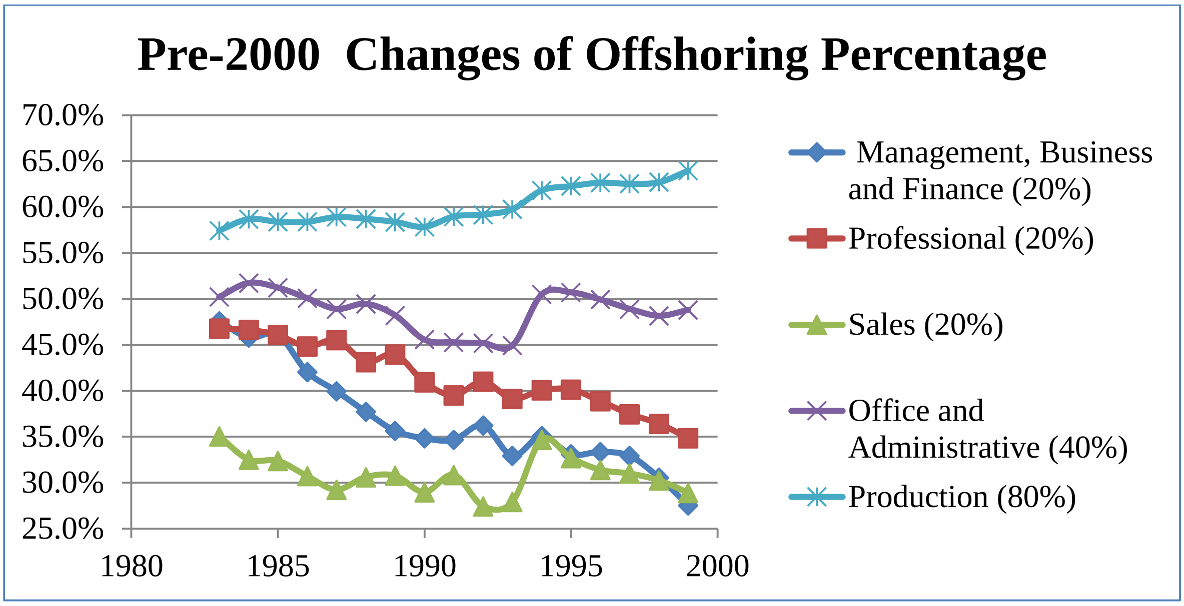
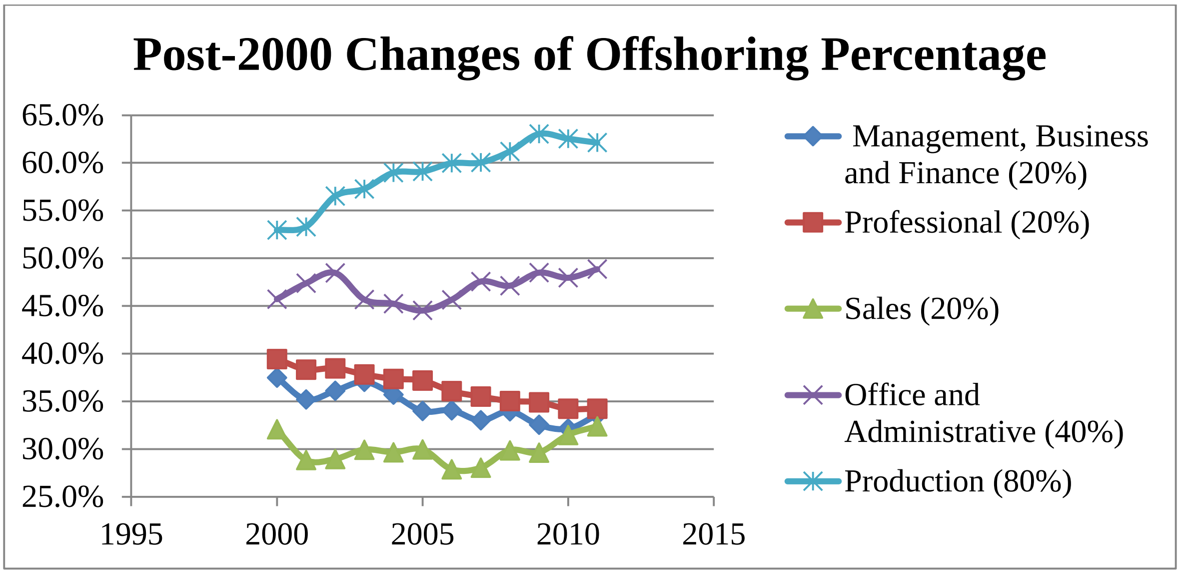


Figure 3.5: Changes of Offshoring Percentage for the Five Relatively Offshorable Occupational Groups





REFERENCES

REFERENCES

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics 118* (4), 1279–1333.

Bardhan, A. D. and C. Kroll (2003). The new wave of outsourcing. Technical report, UC Berkeley: Fisher Center for Real Estate and Urban Economics.

Blinder, A. S. (2006). Offshoring: The next industrial revolution. *Foreign Affairs 85*, 1113-298.

Blinder, A. S. (2009). How many U.S. jobs might be offshorable. *World Economics 10* (2), 41–78.

Blinder, A. S. and A. B. Krueger (2013). Alternative measures of offshorability: A survey approach. *Journal of Labor Economics 31* (2), S97–S128.

Crinó, R. (2010). Employment effects of service offshoring: Evidence from matched firms. *Economics Letters 107* (2), 253–256.

Criscuolo, C. and L. Garicano (2010). Offshoring and wage inequality: Using occupational licensing as a shifter of offshoring costs. *American Economic Review 100* (2), 439–43.

Ebenstein, A., A. Harrison, M. McMillan, and S. Phillips (2009). Why are American workers getting poorer? estimating the impact of trade and offshoring using the CPS. Working Paper 15107, National Bureau of Economic Research.

Friedman, T. L. (2007). The World Is Flat: A Brief History of the Twenty-First Century. D&M Publishers Incorporated.

Goos, M., A. Manning, and A. Salomons (2008). Recent changes in the European employment structure: the roles of technology, globalization and institutions. Technical report.

Grossman, G. M. and E. Rossi-Hansberg (2008). Trading tasks: A simple theory of offshoring. *American Economic Review 98* (5), 1978–97.

Harrison, A. and M. McMillan (2011). Offshoring jobs? multinationals and U.S. manufacturing employment. *Review of Economics and Statistics 93* (3), 857–875.

Wolberg, G. and I. Alfy (1999). Monotonic cubic spline interpolation. *Proceedings of Computer Graphics International*, 188–195.

Wolberg, G. and I. Alfy (2002). An energy-minimization framework for monotonic cubic spline interpolation. *Journal of Computational and Applied Mathematics 143 (2)*, 145 – 188.