THREE PAPERS IN REGIONAL ECONOMICS: ENERGY PRODUCTIVITY CONVERGENCE, WATER RESOURCE PLANNING, AND WORKFORCE OCCUPATION-INDUSTRY DYNAMICS

BY

JUN WAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Regional Planning
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Geoffrey J.D. Hewings, Chair
Professor Edward Feser
Assistant Professor Katherine Baylis
Yu-Feng Lin (Ph.D), Illinois State Water Survey

ABSTRACT

This dissertation research consists of three papers that use the perspectives of regional economics to examine three key engines of economic growth -- *technology*, *natural resource* and *human capital*, by taking spatial heterogeneities and relationships into consideration. The first paper empirically tests the trade-facilitated technology spillovers in the convergence of energy productivities at the disaggregated manufacturing sectors across European Union (EU) countries. The second paper develops a multi-objective non-linear optimization model to simulate the tradeoffs between streamflow restoration and economic welfare loss in a Chicago suburban county - McHenry County. The third paper establishes a dynamic modeling framework to explore the occupation-industry linkages and decipher an array of labor market signals.

The first paper (chapter 2) differs from most previous empirical convergence studies in the economic growth literature by considering a relatively high degree of sectoral detail - 10 NACE (classification of economic activities in the European community) manufacturing sectors. To account for potential spatial heterogeneity and dependence in regional growth, this chapter adopts a spatial version of energy productivity equation, and extends the single equation cross-sectional setting to time series observations of cross-sectional setting. It reformulates the spatial convergence regression equations into a spatial panel data model with individual (country & industry) effects and uses the panel data procedure for estimation. Under the β-convergence spatial panel-data approach, the estimated coefficients of β are almost all negative and statistically significant. This shows EU countries with low starting energy productivity witness relatively faster growth of energy productivity, and former Eastern-bloc EU countries are catching up to more advanced economies in energy efficiency levels. The estimation results present evidence that energy productivity convergence is conditional upon the cross-country differences in steady-state characteristics. The findings also provide insightful implications for energy and trade policies. It helps project trends in energy productivity and can be used to estimate the sector-specific emission under the European Union's Emissions Trading Scheme (EU ETS). It also affirms that trade does appear to aid technology transfer and therefore promoting trade between

advanced countries and less developed countries becomes necessary.

The second paper (chapter 3) intends to create planning strategies to promote sustainable economic growth while protecting the natural environment by minimizing the occurrence of low streamflows. Steams and associated biological communities are among our most valuable natural resources. Humans rely on the environmental services provided by streams in a myriad of ways. However, in some areas, excessive groundwater pumping exacerbates the already critical pressure on streamflow and needs to be managed through effective planning. Based upon economic and hydrogeological concepts, this paper calculates the amount of streamflow depletion due to groundwater pumping and estimates the negative impact on the socio-economic system if groundwater pumping has to be constrained to restore streamflow. An evolutionary algorithm is used to solve the optimization model and to identify the tradeoff curve (or Pareto frontier) between economic welfare loss and stream flow depletion. The multi-objective optimization is conducted at both county and municipality levels. Comparing municipal Pareto frontiers shows us spatially heterogeneous costs of preserving streamflow through various "shadow prices" and also the different capacity of restoring streamflow. It discusses the shapes of the Pareto frontier, the sensitivity of the pumping boundary constraints, and the sensitivity of return flow coefficients. It concludes that the multi-objective optimization model provides a useful framework to consider conflicting objectives in a typical environmental management and planning process, and that the findings can help decision makers and planners in formulating effective groundwater pumping strategies.

The third paper (chapter 4) seeks answers toward regional labor market dynamics, such as industry-based or occupation-based growth, the sensitivity of occupational demand according to industrial performance, and well-connected industries that show higher multiplier effects in generating jobs. The increasing integration of the world economy has demonstrated the critical need to identify industries and an associated skilled workforce that could help regions maintain their competitiveness. The challenge is to align human capital with current and emerging trends of the regional economy. Bi-causative analysis and hypothetical industrial extraction method are adapted to serve these purposes, with applications to both state and national levels in the U.S. between 2005 and 2008. The findings derived from these linkage studies offers insights on spatially heterogeneous distribution of occupation-based growth, and help identify key industries that cast major demand for each occupation and well-connected industrial sectors that have stronger multiplier effects to generate job growth. Careful monitoring of these signals by manpower planners may provide a means of identifying trends in the balance of skills demand that can be used to detect structural changes in the regional economy and guide manpower planning practice. It provides a basis for determining the desirable level of public and private expenditure on specific education and training programs, and necessary assistance to industries.

To mom and dad - the support and love you provide to me are tremendous

ACKNOWLEDGEMENTS

I would like to express my gratitude to the many people who provided me with assistance, comments, and suggestions as I pursued this work. First and foremost, I am deeply indebted to Professor Geoffrey J. D. Hewings for his patience, guidance, and financial support to make this dissertation possible. I am inspired by the high standards that he sets for himself and those around him, his attention to the details, his dedication to the regional science field, and his passion about work and life. I also want to express my gratitude for my committee members, Professor Edward Feser, Professor Katherine Baylis, and Dr. Yu-Feng Lin. Their insightful suggestions and comments made toward my previous work helped me finally come up with this dissertation.

Working in the Regional Economics Applications Laboratory (REAL) has been a very precious time in my life. I am so blessed to have such a wonderful group of people who create a splendid environment for sharing ideas and working together toward a common goal. I will miss the discussions and snack time in these numerous late nights, and our REAL *mafia* fun time at the Murphy's Pub which we call the alternative office.

I would like to give special thanks to Dr. Ethan Yang, Jihua Wang, and the staff scientists from the Illinois State Water Survey who offered me invaluable interdisciplinary collaboration opportunities and research resources on water resource planning project. Their detailed explanations of hydrologic concepts and the provision of multi-objective genetic algorithm code were particularly helpful. I would also like to thank Professor Peter Mulder from VU University Amsterdam for providing me with the data on energy use for detailed industrial sectors across European Union countries. Without them, this dissertation could not even be close to finishing.

Thanks also go to my previous advisor Professor Bill Anderson at University of Winsor who introduced me to the field of regional science through an independent study on economic geography. I still can clearly remember the discussions we had years ago and his encouragement and help for me to pursue a doctoral degree in Regional Planning at University of Illinois.

Finally, I owe a considerable debt of gratitude to my parents in China who made all of this possible with their endless support and understanding. I can't imagine how much they have suffered over the past years

when they have to eat at the dinner table without my presence. Their words helped me overcome the depression and frustration I encountered during my graduate studies. Their love and unyielding patience have enriched my life and I am so lucky to always have them available at the other end of the telephone line. My heart will be forever yours.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

During the past a few decades, economists have spared no effort to look for the engines of economic growth. In general, the most frequently visited three topics are *technology*, *natural resource* and *human capital*. One novelty of regional economics is to incorporate space and spatial relations into traditional economic studies. Recognition of uneven geographic distribution of technology, natural resource and human capital helps reconsider their role in promoting regional growth. Studies from a regional economics' perspective could reveal the evolving spatial interdependences of economic relations when growth occurs. It can provide more insights than traditional economic analysis to guide policy makers and regional planners. This dissertation intends to explore economic activities associated with those three growth engines by taking spatial heterogeneities and relationships into consideration.

In chapter 2, the main purpose is to ascertain possible trade-facilitated spatial correlation in the energy productivity improvement process across EU countries. Energy productivity is defined as output divided by final energy use, and in manufacturing sectors it highly depends on the technological advancement of the production process. Countries lagging behind in terms of energy productivity levels may benefit from experiences and technologies developed by countries operating at the forefront, a process that might lead to convergence of cross-country energy productivity levels. Also, spatial correlations associated with trade flow relations might contribute to the technology catch-up due to knowledge spillovers and technology diffusion. A spatial panel model is hence constructed to test two relevant concepts: convergence and spatial technology spillovers. The findings could provide insightful implications for energy and trade policies. It helps project trends in energy productivity and can be used to estimate the sector-specific emission under the European Union's Emissions Trading Scheme (EU ETS). It also tests the hypothesis whether trade would aid or promote technology transfer.

Chapter 3 proposes an integrated hydrologic-economic framework to study groundwater resources under the context of a fast urbanization process. Excessive groundwater withdrawal can reduce environmental amenities by depleting the stream flow since wells and nearby streams are hydraulically connected to the same aquifer. Constraining well pumpage to return streamflow may result in consumer welfare loss. Therefore, effective and holistic management of groundwater resources becomes imperative, if it could incorporate concerns of preserving environmental amenities while minimizing the negative impacts on social, economic and political systems. In this chapter, tradeoffs between environmental amenity and economic welfare loss from adopting spatially targeted pumping limits will be discussed and analyzed. The planning decisions will be based on the localities of wells and streams, physical hydrologic facts, and groundwater-supported socio-economic activities.

The increasing integration of the world economy has demonstrated the critical need to identify industries and an associated skilled workforce that could help regions maintain their competitiveness. The challenge is to align human capital with current and emerging trends of the regional economy. Chapter 4 establishes an empirical modeling framework to explore the occupation-industry linkages and decipher an array of labor market signals, such as industry-based or occupation-based growth, the sensitivity of occupational demand according to industrial performance, and well-connected industries that show higher multiplier effects in generating jobs. The findings derived from these linkage studies offer insights on spatially heterogeneous distribution of occupation-based growth, and help identify key industries that cast major demand for each occupation and well-connected industrial sectors that have stronger multiplier effects to generate job growth. Careful monitoring of these signals by manpower planners may provide a means of identifying trends in the balance of skills demand that can be used to detect structural changes in the regional economy and guide manpower planning practice. It provides a basis for determining the desirable level of public and private expenditure on specific education and training programs, and necessary assistance to industries.

Chapter 2

TESTING OF TRADE-FACILITATED TECHNOLOGY SPILLOVERS IN ENERGY PRODUCTIVITY CONVERGENCE PROCESS AMONG EUROPEAN COUNTRIES

2.1 Introduction and Literature Review

Given the threat of disruptive climate change and high energy prices in recent years, the ability of nations to reduce their reliance on fossil fuels and subsequently lower their emission level of greenhouse gases becomes a primary concern of scientists, economists, and politicians. Increasing energy productivity or the economic output associated with a given unit of energy use is considered as an important policy objective to further economic development while reducing the risks of global warming (Patterson, 1996). Energy productivity is defined as output divided by final energy use. The world could reduce the growth of energy demand by raising energy productivity, and improved energy productivity level could also help assuage concerns about how to secure future energy supplies.

Energy productivity in manufacturing sectors highly depends on the technology used in the production process. Traditional growth models simply assume all regions having access to the same blueprint of technology (Borts and Stein, 1964). We know, however, that spatial differences in technology do exist. Given that, are cross-country differences in energy productivity performance decreasing or is the gap between the leading and "backward" countries widening? Does geographical proximity or other forms of connectivity facilitate diffusion of energy saving technologies? To answer these questions, this chapter examines two concepts related to energy productivity: convergence and technology diffusion.

Convergence in this context is assumed to involve the decrease in cross-country differences in energy productivity levels. The rationale behind this hypothesis is that countries lagging behind in terms of energy productivity levels can benefit from the experiences and technologies developed by countries operating at the forefront, a process that might lead to convergence of cross-country energy productivity levels. The concept of productivity convergence can be traced back to the traditional Solow-Swan neoclassical growth models with its central notion of a transitional growth path toward a steady state, due to diminishing returns to capital accumulation (Solow, 1956; Swan, 1956). Assumptions and predictions from theoretical convergence models have been tested by a large number of empirical contributions focusing on regional productivity growth and cross-country convergence of per capita income (Barro, 1991; Barro and Sala-i-Martin, 1992; Mankiw *et al.*, 1992; Sala-i-Martin, 1996; Carlino and Mills, 1996;

Bernard and Jones, 1996). The modern or endogenous growth theory (Lucas, 1988; Romer, 1986) produces a more diverse picture towards patterns of convergence, by considering knowledge accumulation and diffusion of technology. Following the seminal work by Coe and Helpman (1995), a growing trend in the literature has treated diffusion of technology across countries as a source of promoting productivity and thus eliminates the tendency toward diminishing returns prevalent in neoclassical models.

Given that energy is a principal input in most of the production functions (Hudson and Jorgenson, 1974; Berndt and Wood, 1975) in addition to traditional factors such as labor and capital, similar concepts of convergence and catch-up have been applied to the field of energy productivity or energy intensity developments (Miketa and Mulder, 2005; Markandya *et al.*, 2006; Mulder and Groot, 2007). Results provide support for the hypothesis that in most sectors lagging countries tend to catch up with technological leaders, particularly in terms of energy productivity. Figure 2.1 shows the energy productivity (1997 US dollar per kilo tons of oil equivalent) for 8 European countries in the manufacturing sector from 1995 to 2005. It is easy to notice that countries starting with low energy productivity levels (e.g., Slovakia, Poland, Czech Republic) are catching up with energy more efficient countries (e.g., Netherlands and Denmark). And the variances of energy productivity among these countries are decreasing across years, also indicating a converging process.

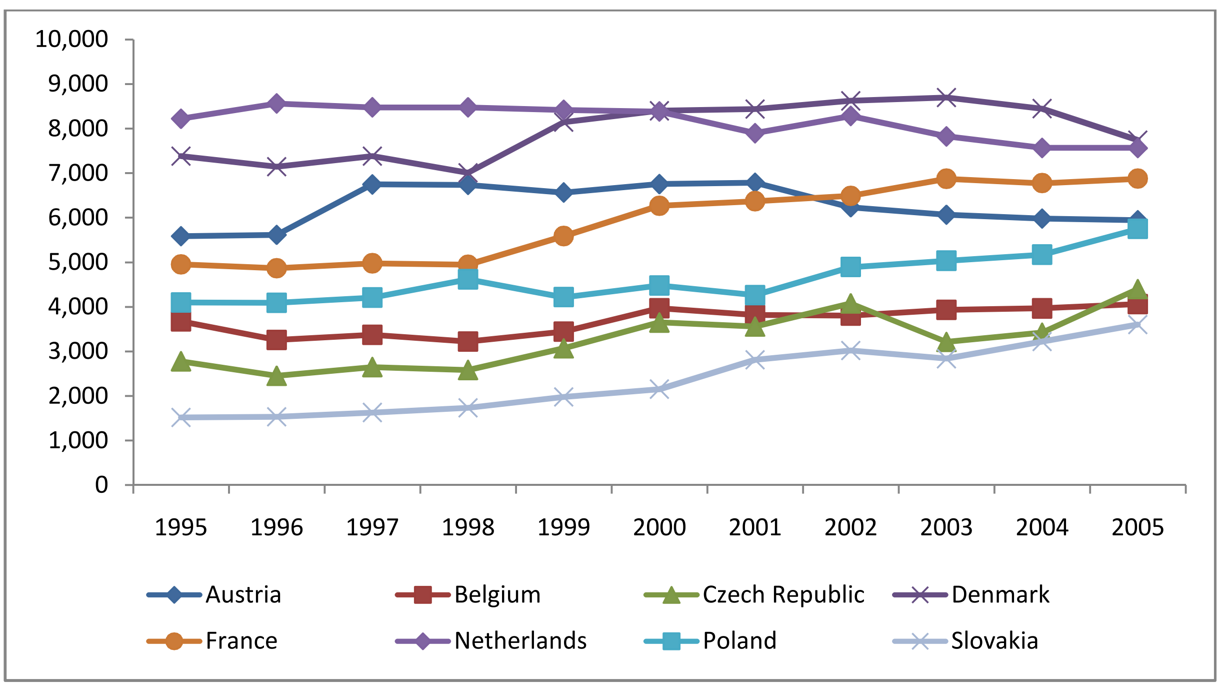

Figure 2.1: Energy Productivity Dynamics for 8 EU Countries in Manufacturing Sector (1995 - 2005)

This chapter considers a relatively high degree of sectoral detail. This is important because analysis of energy productivity convergence at an aggregated level of industrial sector would mask internal structural change and substantial differences in energy productivity convergence dynamics at the sectoral level. Also it is well noticed that technologies are likely to be more transferrable within a sector than among sectors.

This chapter addresses the problem of previous analyses on energy productivity convergence that have ignored the fact that processes of technology catch-up might be the result of technology diffusion. Technology spillovers could go beyond national (geographical) boundaries. International technology diffusion has been an important factor to determine the speed at which the world's technology frontier expands. The idea that trade might be contributing to the international transmission of technology has been frequently emphasized by scholars. For instance, Coe and Helpman (1995) presented evidence to show that technology spills over across countries through the channel of trade flows, and provided estimates of the magnitude of these spillovers.

More specifically, in the context of energy productivity, export orientation could be a central determinant of energy efficiency innovation (Urpelainen, 2011). Current state-of-the-art energy efficient technologies created by concerted research efforts are embedded in the commodities, and the knowledge captured by the inventions could spill over to the destinations through bilateral trade linkages. From a macro-perspective, domestic industry sectors tend to undergo a self-selecting process into export markets with high productivity, including high energy productivity. Strong trade linkages (both intra- and inter-industry trade) then help increase the absorptive capacity (Cohan and Levinthal, 1990) and structural similarity (Hayami and Ruttan, 1985) of energy efficient technologies between the trading nations. From a micro-perspective, in order to maintain competitive advantages, firms have strong incentives to develop and commercialize new energy technologies, especially when facing rocketing energy prices. Usually, intra-industry trade is not confined to the trade of final products but also includes the trade of R&D results and technologies, especially energy efficient technologies, innovations used during the production process (e.g. patents), technology embodied intermediate goods, and etc. For instance, Siemens has made energy efficiency innovation a key part of its business strategy, and they are now providing equipment, patents, and technology through trade to a variety of countries across the global, such as wind turbine technology.

This chapter differs from previous work (Coe and Helpman, 1995; Coe and Helpman, 1995; Park, 1995) by using sector-level trade transaction data to proxy for the intensity of technology spillovers, as opposed to the country-level trade transaction data. The rationale behind is that studies employing aggregate data are likely to miss the intra-industry dynamics of the technology flows by ignoring the diversity of sectoral characteristics. The use of two- or three-digit industry level trade transaction data should reduce this problem. Through identifying the trade flow relations at detailed sectoral level among these EU countries, it helps draw a clear picture on which partner countries are closely related and further helps trace the possible strong technology spillovers in the energy productivity improvement process. To accommodate the trade-facilitated spillover effects in energy-productivity growth specifications, this chapter harnesses the spatial econometric techniques (Amstrong, 1995; Rey and Montouri, 1999; Fingleton and McCombie, 1998, Fingleton, 1999, Lopez-Bazo *et al.* 1999) by specifically considering the trade flow weight matrix.

The following section presents the methodology, introducing the structural model on convergence, its specification, the selection of weight matrix, and the panel specification of convergence models. Section 2.3 reviews the sources of data used, and section 2.4 reports and interprets the empirical analytical results of the spatial panel energy productivity convergence models. A concluding discussion is provided in section 2.5.

2.2 Methodology

2.2.1 A Structural Growth Model on Energy Productivity Convergence with Technology Spillovers

To illustrate the concept of the various factors and mechanisms that may affect cross-country energy productivity differences and evolution dynamics, this chapter adopts a neoclassical Cobb-Douglas production function that uses labor (L), capital (K), and energy (E) to produce its output (Y):

$$Y_{it} = A_{it} K_{it}^{\alpha} L_{it}^{\beta} E_{it}^{\gamma} \qquad (2\text{-}1)$$

A_{it} denotes the technology at country i in period t. And since constant returns to scale are assumed in this formulation, I have further:

$$\gamma = 1 - \alpha - \beta \qquad (2\text{-}2)$$

Average labor productivity in country i in period t, y_{it}, is a function of capital and energy per unit of labor (k_{it} and e_{it}):

$$y_{it} = A_{it} k_{it}^{\alpha} e_{it}^{\gamma} \qquad (2\text{-}3)$$

In terms of energy productivity (Y_{it}/E_{it}), equation (2-3) can be written as:

$$\frac{Y_{it}}{E_{it}} = A_{it}^{\frac{1}{\gamma}} \left[\frac{K_{it}}{L_{it}}\right]^{\frac{\alpha}{\gamma}} \left[\frac{Y_{it}}{L_{it}}\right]^{\frac{\gamma-1}{\gamma}} = A'_{it} k_{it}^{\frac{\alpha}{\gamma}} y_{it}^{\frac{\gamma-1}{\gamma}} \qquad (2\text{-}4)$$

Similar to the setup of technological externalities adopted by Lopez-Bazo *et al*. (2004), A'_{it} $(= A_{it}^{\frac{1}{\gamma}})$ in the equation (2-3) is assumed to depend on the technological level of trading partners or neighboring countries:

$$A'_{it} = \Delta_t (k_{\vartheta it}^{\alpha} e_{\vartheta it}^{\gamma})^m \tag{2-5}$$

where Δ_t is an exogenous component assumed identical for all economies, with a growth rate equal to g ($\Delta_t = \Delta_0 e^{gt}$). Expressions $k_{\vartheta it}^{\alpha}$ and $e_{\vartheta it}^{\gamma}$ denote the capital- and energy-labor ratios in the trading partners or neighboring countries. Finally, m measures the externalities across economies which are assumed to be positive. Therefore, under this assumption, new ideas and technological innovations flow or diffuse across economies so that an economy benefits from investment made by its trading partners or neighbors. Combining equation (2-5) with equation (2-3) and (2-4) yields an expression that relates labor productivity and energy productivity in a country to capital intensity in the same country and in its trading partners or neighboring countries:

$$y_{it} = \Delta_t k_{it}^{\alpha} e_{it}^{\gamma} (k_{\vartheta it}^{\alpha} e_{\vartheta it}^{\gamma})^m \tag{2-6}$$

$$\frac{Y_{it}}{E_{it}} = \Delta_t (k_{\vartheta it}^{\alpha} e_{\vartheta it}^{\gamma})^m k_{it}^{\frac{\alpha}{\gamma}} y_{it}^{\frac{\gamma-1}{\gamma}} \tag{2-7}$$

Under the assumption of decreasing returns to capital and energy within each region, $(\alpha + \gamma) < 1$, the expressions for the steady state of capital and energy per effective labor can be written as:

$$\tilde{k}^* = \left(\frac{s_k^{1-\alpha} s_e^{\gamma} \tilde{k}_{\vartheta}^{m\alpha} h_{\vartheta}^{m\gamma}}{n+g+d}\right)^{\frac{1}{\beta}} \tag{2-8}$$

$$\tilde{e}^* = \left(\frac{s_k^{\alpha} s_e^{1-\gamma} \tilde{k}_{\vartheta}^{m\alpha} h_{\vartheta}^{m\gamma}}{n+g+d}\right)^{\frac{1}{\beta}} \tag{2-9}$$

while for output per unit of effective labor will reach a steady state $\tilde{y}^*$:

$$\tilde{y}^* = \left(\frac{s_k^{\alpha} s_e^{\gamma} \tilde{k}_{\vartheta}^{m\alpha} h_{\vartheta}^{m\gamma}}{(n+g+d)^{\alpha+\gamma}}\right)^{\frac{1}{\beta}} \tag{2-10}$$

where s_k and s_e are the rates of accumulation of capital and energy and n, g, and d are population, technology growth, and the rate of depreciation, respectively. The sum $n+g+d$ is what has been called effective rate of depreciation in the growth literature and is assumed to be equal across types of capital and economies. From equations (2-8), (2-9), and (2-10), it can be observed how the technology of production is characterized by externalities across economies. The steady state depends on the usual technological and preference parameters but also on capital intensity in the trading partners or neighboring countries. If the steady states of $\tilde{k}^*$, $\tilde{e}^*$, and $\tilde{y}^*$ are plugged into equation (2-7), the energy productivity will reach a steady state as well. The influence of spillover or technological diffusion effect will be greater, the larger the returns to capital and the coefficient that measures the strength of the externality, m.

2.2.2 Specifications of Convergence

There are two traditional specifications that have been extensively used in the literature to analyze convergence, namely β–convergence and σ–convergence. The β–convergence tests whether a statistically significant negative relationship exists between the initial level and the growth rate of energy productivity. Alternatively, the σ–convergence refers to the method that calculates the un-weighted standard deviation (Kuznets, 1955; Easterlin, 1960a, 1960b; Williamson, 1965; Amos, 1989; Coughlin and Mandelbaum, 1988; Fan and Casetti, 1994; Carolin and Mills, 1996; Bernard and Jones, 1996) of energy productivity levels cross-country over time.

Under the β–convergence specification, considering the fact that energy productivity convergence depends to a large extent on individual country-effects, it is necessary to incorporate additional variables (X) in the specification to control for factors determining steady states across different countries. The resulting specification is of the following form:

$$g_{EP} = c + \beta \ln(EP_0) + X\delta + \varepsilon \tag{2-11}$$

where g_{ep} is the growth rate of EP (energy productivity, defined as $\frac{Y}{E}$) in a country for a given period, EP_0 is its initial energy productivity level, c denotes the intercept, ε is a well-behaved error term and the scalar β is the measure of the speed of convergence. When $\beta < 0$ and significant, and δ is a vector whose elements are non-significant, the energy-productivity is in favor of absolute β-convergence, while in the case of $\beta < 0$ and significant, and δ is a vector of significant coefficients, the outcome is a conditional β-convergence.

2.2.3 Spatial Convergence Specification and Selection of Weight Matrix

To account for potential spatial heterogeneity and dependence in regional growth specifications, this chapter adopts a spatial version (Armstrong, 1995; Rey and Montouri, 1999; Lopez-Bazo *et al.*, 1999; Bivand and Brundstad, 2006) of energy productivity equation. The spatial version of the convergence equation includes the spatial lag of the growth rates (spatial lag model) and a spatial structure in the perturbance (spatial error model).

The choice of spatial weight matrix is mostly based on the researcher's assumption about how regional externalities occur (Li and Haynes, 2011). Frequent options for representing spillover effects would be the contiguity matrix based on adjacency, or a distance matrix based on physical distance. However, in the case of the EU countries, neighboring countries often do not have the same mother tongue and their technological interaction might not be well captured by the pure physical distance. Evidence shows that industries or firms that engage in international trade would be able to raise their productivity by interacting with technologically advanced trading partners (Keller, 2009). This fact could also be true to energy productivity when energy efficient technology diffusion appears to be increasing along with higher levels of economic integration. Therefore, this chapter chooses international trade activities among EU countries at each individual manufacturing sector to produce weight matrix.

After specifying the weight matrix, a spatial lag model of energy productivity convergence can be expressed as:

$$g_{EP} = c + \beta \ln(EP_0) + X\delta + \rho W g_{EP} + \varepsilon \tag{2-12}$$

where $W g_{EP}$, the spatial lag of energy productivity growth rates, is obtained by pre-multiplying the vector of energy productivity growth rates of different countries by the trade-flow-based spatial weights matrix, W. This matrix is a pre-specified, non-negative, row-standardized matrix and describes the total amount of trade flows (including exports and imports) among EU countries. The variable ρ is the spatial autoregressive coefficient and $\varepsilon \sim N(0, \sigma^2)$.

The expression for the spatial error model about energy productivity convergence is:

$$g_{EP} = c + \beta \ln(EP_0) + X\delta + \varepsilon, \qquad \varepsilon = \lambda W \varepsilon + \tau \tag{2-13}$$

$$g_{EP} = c + \beta \ln(EP_0) + X\delta + (I - \lambda W)^{-1} \tau \tag{2-14}$$

In this case, a random shock in a country affects energy productivity growth rates in that country, and additionally impacts all other trade-related countries through the spatial transformation. The remaining disturbance term follows a first-order serially auto-correlated process with $e_t \sim N(0, \sigma_e^2)$. Variable λ ($|\lambda| < 1$) is the spatial autocorrelation coefficient and φ ($|\varphi| < 1$) is the serial autocorrelation coefficient. The model allows for serial correlation on each spatial unit over time as well as spatial dependence between spatial units at each time period.

$$\tau_t = \varphi \, \tau_{t-1} + e_t \tag{2-15}$$

2.2.4 A Panel Data Approach toward Energy Productivity Convergence

Most of the studies on energy productivity convergence are conducted under the framework of single cross-country regressions, by assuming identical production function for all the countries in the specified industrial sectors. However, the panel data framework (Islam, 1995) makes it possible to correct this bias by allowing for differences of the above-mentioned type in the form of unobserved individual "country effects" and "industry effects". It extends the single equation cross-sectional setting to time series observations of cross-sectional setting. The use of panel data offers a greater availability of degrees of freedom, increases the efficiency in the estimation, and avoids the collinearity among the variables (Elhorst, 2003). This chapter thus reformulates the spatial convergence regression equations into a spatial panel data model with individual (country & industry) effects and uses the panel data procedure to estimate it.

More specifically, under the panel data structure, the spatial lag and error models of energy productivity convergence can be re-written as:

$$g_{EP_{ij,t}} = c + \beta \ln\left(EP_{0_{ij,t}}\right) + X_{ij,t}\delta + \rho W g_{EP_{ij,t}} + \varepsilon_{ij,t} \tag{2-16}$$

$$g_{EP_{ij,t}} = c + \beta \ln\left(EP_{0_{ij,t}}\right) + X_{ij,t}\delta + \varepsilon_{ij,t}, \qquad \varepsilon_{ij,t} = \lambda W \varepsilon_{ij,t} + \tau_{ij,t} \tag{2-17}$$

where $EP_{0_{ij,t}}$, $g_{EP_{ij,t}}$ are starting year's and the growth of energy productivity for country *i* and industry *j* at the *t*th time period, and $X_{ij,t}$ are the control variables for country *i* and industry *j* at the *t*th time period. Trade-flow-based spatial weight matrix (*W*) for convergence models of panel structure is compiled as:

$$W = \begin{bmatrix} TF_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & TF_j \end{bmatrix}, TF_n = Row\ Normalize\ [F_n], \text{ and } F_n = \begin{bmatrix} 0 & tf_{21} & \cdots & tf_{i1} \\ tf_{12} & 0 & \cdots & tf_{i2} \\ \vdots & \vdots & 0 & \vdots \\ tf_{1i} & tf_{2i} & \cdots & 0 \end{bmatrix} \tag{2-18}$$

$$tf_{pq} = tf_{qp} = Export(p,q) + Import(p,q) \tag{2-19}$$

where $p, q = 1, 2, \dots, i$ and $n = 1, 2, \dots, j$. The diagonal of spatial weight matrix (*W*) is composed of row normalized trade flow matrices ($TF_n, n = 1, 2, \dots, j$) for each detailed manufacturing sector specified in this chapter, and the off-diagonal values are set to zero. The elements (tf_{pq}) of trade flow matrix $[F_n]$ is formed by the nominal dollar values of trade flow (exports plus imports) at a given year for sector *n* ($n = 1, 2, \dots, j$) between country p and country q. The diagonal values of matrix $[F_n]$ are set to zero, indicating that each country does not trade with itself (no sub-country regional trade flow is assumed in this case).

2.3 Data

The analysis presented in this chapter is based on a newly constructed database that combines the newly launched "EU KLEMS Growth and Productivity Accounts" with physical energy data from the International Energy Agency (IEA). Economic and trade flow data are from the International Sectoral Database (ISDB) and the Structural Analysis Database (STAN), both published by Organizations for Economic Co-operation and Development (OECD). Construction of this database helps establish a link between economic and energy data at a detailed sectoral level. Table 2.1 shows the sector classification. This enables us to do a systematic cross-country convergence analysis of energy productivity performance at a high level of sectoral detail. The database covers the period 1970 - 2005 and includes the following 16 EU countries (see figure 2.2): Austria (AUT), Belgium (BEL), Czech Republic (CZE), Denmark (DNK), Finland (FIN), France (FRA), Germany (GER), Hungary (HUN), Italy (ITA), Netherlands (NLD), Poland (POL), Portugal (PRT), Spain (SPA), Slovakia (SVK), Sweden (SWE), and United Kingdom (UK).

Table 2.1: Sector Classification

	Sector	Abbreviation
1	Food, Beverages and Tobacco	FOD
2	Textiles, Leather and Footwear	TEX
3	Wood and of Wood and Cork	WOD
4	Pulp, Paper, Printing and Publishing	PAP
5	Chemicals	CHE
6	Non-Metallic Minerals	NMM
7	Basic Metals	BMI
8	Machinery	MAC
9	Transport Equipment	TRE
10	Non-Specified Industry	NSI

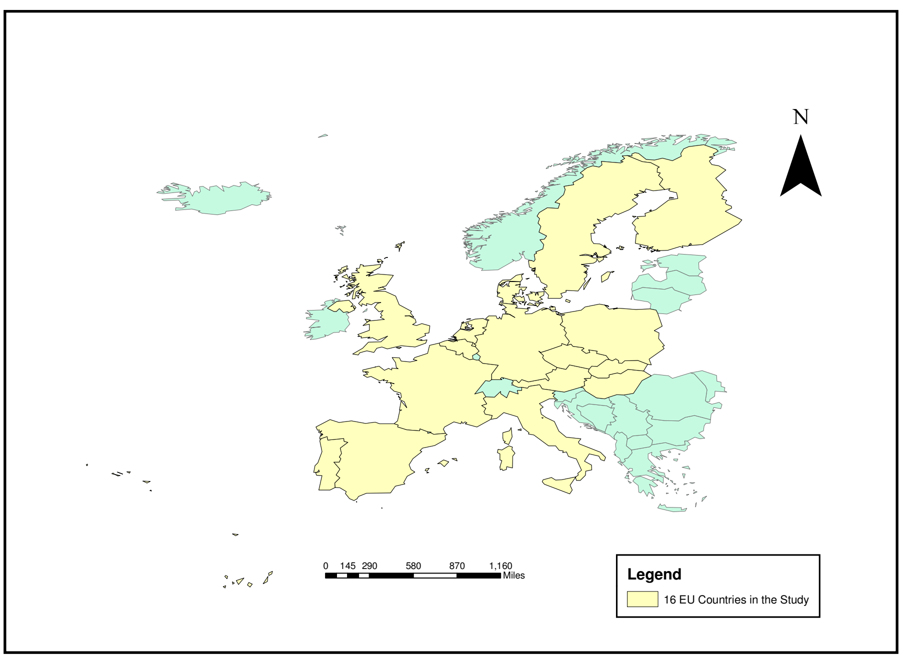

Figure 2.2: Map of European Countries

Energy productivity is measured by gross value added per unit of final energy consumption. Value added is the net economic output of a sector, measured by the price differential between the price of output and the cost of input. Value added data have been converted to constant 1997 US $, using industry-specific Purchasing Power Parities (PPPs) for 1997 (Mulder and Groot, 2011). Moreover, this database includes data on capital compensation (*CAP*), gross output (*GOS*). According to EU KLEM, energy use is defined as final energy consumption in kilo tonnes of oil equivalent (ktoe), with sectoral data excluding transformation losses. Sector-specific constant energy prices (1997 US$/ktoe) are also based on EU KLEM data.

Trade flow data (including both imports and exports) for each industry sector among the 16 EU partner countries are retrieved from 2006 STAN Bilateral Trade Database, and the latest year's data (2004) are taken to estimate the benchmark level of trade relations among the EU countries at detailed sectoral level. The unit of measure is in thousands US $ at current price.

2.4 Results of Spatial Panel Convergence Regressions

2.4.1 Selection of Conditional and Instrument Variables

This chapter searches for country-specific sectoral determinants (*X*) of energy productivity growth by including a number of country-specific explanatory variables in the spatial panel convergence regression models.

Heating Degree Days (*HDDs*) and Cooling Degree Days (*CDDs*): A "degree day" is a measure of the average temperature's departure from a human comfortable level of 18 °C (65 °F). The concept of degree days is used primarily to evaluate energy demand for heating and cooling services. *HDDs* and *CDDs* are both calculated in a cumulative fashion. Data on *HDDs* and *CDDs* are collected from the Climate Analysis Indictors Tool (CAIT) at World Resource Institute (Baumert and Selman, 2003).

Capital Share (*K_SHARE*): Capital Share is a measure of the capital use intensity in manufacturing production process, and is calculated by the ratio of capital compensation (*CAP*) to gross output (*GOS*). The accumulation of capital stock is associated with sector-specific learning-by-doing and high capital share raises the amount of sectoral public knowledge. It serves as a good indicator of country- and industry-specific level of development.

Energy Price (*EPRIC*): Price information on energy use for each country is calculated as sector-specific energy expenditure (in constant 1997 US $) divided by physical energy consumption (in ktoe). Energy expenditure (in $) and physical energy consumption (in ktoe) are both retrieved from EU KLEM database, where energy is defined as an intermediate input derived from national accounts and supply-and-use tables. However, in most cases, the price of energy is correlated with the energy productivity series that influence the growth rate of energy productivity via the convergence model specification, and in turn the estimation results from the regression are biased. A Two-Stage Least Square (2SLS) estimation is then adopted to address the issue of endogeneity. In the first stage, energy prices are regressed on a series of variables that do not endogenously correlate with the energy productivities (see equation 2-19). For the second stage, the spatial panel convergence regression models (see equations 2-16 and 2-17) then use the predicted energy prices from the first stage as an independent variable. Instrumental variables are selected to correct for the corresponding bias by replacing the original energy price variable with the new variable that is uncorrelated with the disturbance term of the dependent variable - energy productivity. In other words, the fitted values of $\log{(EPRIC)_{ij,t}}$ of step 1 are used for a maximum likelihood estimation applied to the stacked panel structural model.

$$\log{(EPRIC)_{ij,t}} = \alpha_0 + \alpha_1\, PCP_{OIL_{ij,t}} + \alpha_2\, PCP_{GAS_{ij,t}} + \alpha_3\, PCP_{COAL_{ij,t}} + \alpha_4 \log(CDDs + HDDs)_{ij,t} + \alpha_5\, K\_SHARE_{\ ij,t} \quad (2\text{-}19)$$

In this chapter, the instrumental variables include per capita oil reserve (PCP_{OIL}), per capita natural gas reserve (PCP_{GAS}), per capita coal reserve (PCP_{COAL}), logarithm of the sum of *HDDs* and *CDDs*, and the *K_SHARE*. Data on international energy reserves and resources are obtained from U.S. Energy Information Administration (EIA). These variables serve as good instruments for the price variable, mainly because the natural endowment of energy resources, weather conditions, and the capital share of production process are correlated with the energy price but meanwhile they are all exogenous to and do not depend on the energy productivity.

LEAD: Grams of lead content per gallon of gasoline are used as a proxy for the level of environmental stringency at the country level. Given that lead emissions are precursors to harmful local air pollutants, a country with relatively strict environmental policy should allow lower lead content per gallon of gasoline. *LEAD* data are collected from "worldwide gasoline survey" published annually by OCTEL (1983-1995), and this chapter uses the data from year 1995.

Also, five types of dummy variables are created. *INDUSTRY* is a manufacturing sector associated dummy variable, to test the varying effects associated with each sector; *COUNTRY* is a country-specific dummy variable, to test the varying effects associated with each EU country; *TP* is a time period dummy variable, to specify the time associated effects; *NE* is a dummy variable for "non-eastern" European regions; *H_EPRIC* is a dummy variable for top 4 countries (25th percentile) with the highest energy price for each specified manufacturing sector at a given testing time period. In this chapter, Czech Republic (CZE), Hungary (HUN), Poland (POL), and Slovakia (SVK) are classified as eastern European countries, and the rest 12 countries are considered as non-eastern European countries.

2.4.2 Selection of Estimation Periods of Convergence

Given that current spatial statistical package can only estimate spatial panel regression with balanced panel dataset, the selection of testing periods for convergence among specified EU countries is confined to the period of 1995-2005. Even though, such a time period under study still covers a sufficiently long horizon. Considering that the underlying dataset on cross-country & cross-industry energy productivity are available at the annual level, however, it seems that yearly time spans are too short to be appropriate for studying growth convergence. Instead, this chapter opts for five-year intervals (1995-2000, 2000-2005) to minimize business cycle effects. Therefore, the panel data formulation is able to turn a single cross-section for the entire period (1995-2005) to cross-section for shorter periods that constitute it. In other words, after controlling for the individual country effects, the panel data setup helps to integrate the process of convergence occurring over several consecutive time intervals.

There has been a variety of energy policies applied to the manufacturing sector. During 1990s, voluntary agreements for energy efficiency improvement and reduction of energy-related GHG emissions by industry were dominant (Price, 2005). First agreements were the Long-Term Agreements on Energy Efficiency in the Netherlands, the Danish Agreement on Industrial Energy Efficiency, and the Declaration of German Industry on Global Warming Prevention. Since year 2000, a number of countries that first established strictly voluntary agreements started to strength their programs in a follow-on phase. For instance, in 2002, France replaced its initial 1996 voluntary program with a new program that includes a penalty fee for non-compliance and allows for emission trading. Other countries that have a second generation of agreements, such as Finland, Germany, and The Netherlands, all either increased the number of incentives or added penalties to strengthen the programs (Price, 2005). Therefore, it is worthwhile to test the policy effects on energy productivity growth rates for the two periods (1995-2000 and 2000-2005).

2.4.3 Analysis of Results

In the panel regression models (equations 2-16 and 2-17), it is very likely that the error term contains all sorts of (unobserved) country- & industry-specific tangible and intangible factors that affect energy productivity growth. One common issue that arises in such estimation is whether the individual effects are to be thought of as "fixed" or "random". Typically, panel approach applies fixed- or random-effects models to solve this problem, and the random-effects model assumes the effects are uncorrelated with the exogenous variables included in the model (Islam, 1995). The presence of the spatial lag introduces a form of endogeneity that violates the assumption of the standard regression model that the regressors are uncorrelated to the error term. One frequently used estimation method for spatial models is maximum likelihood (ML) which then accounts for endogeneity. However, the ML estimation entails substantial computational problems if the number of cross sectional units is large (Kapoor, Kelejian, and Prucha, 2007). To solve this problem, a generalized moments (GM) estimation was proposed by Kelejian and Prucha (1999). This chapter adopts the ML estimation methods based on the fact that there is a manageable size of cross sectional units (160 units) for the panel.

First, equations (2-16) and (2-17) are estimated by using the fixed effects (FE) estimator. More specifically, for the FE specification, following Elhorst (2003), these models can be specified with "pooled effects" (only the optional constant term is included), "time period specific effects", and "both spatial and time period fixed effects". The results of "pooled effects" for spatial lag model and spatial error model are reported in table 2.2 and table 2.3.

Table 2.2: Estimation Results of Fixed Effect Spatial Lag Models under "Pooled Effects"

	Case 1		Case 2		Case 3		Case 4		Case 5		Case 6		Case 7	
Intercept	-3.446	***	-3.464	***	-1.611	*	-3.562	***	-3.583	***	-0.945		0.126	
log(EP_0)	-0.030	**	-0.025	*	-0.027	*	-0.032		-0.011		-0.392	***	-0.160	**
EPRIC	0.112	***	0.047		0.053	*	0.121	***	0.113	**	0.184	***	0.111	**
log(HDDs + CDDs)	0.300	***	0.390		0.188	***	0.308	***	0.297	***	0.257	***		
K_SHARE	-1.195	***	-1.403	***	-1.403	***	-1.189	***	-1.254	***	-1.332	***	-1.587	***
LEAD			1.771	***										
NE					-0.200	***								
log(EP_0): Dummy(T2)							0.006							
log(EP_0): Dummy(H_EPRIC)									-0.078	**				
log(EP_0): Dummy(TEX)											0.090			
log(EP_0): Dummy(WOD)											0.304	**		
log(EP_0): Dummy(PAP)											0.219			
log(EP_0): Dummy(CHE)											0.389	**		
log(EP_0): Dummy(NMM)											0.075			
log(EP_0): Dummy(BMI)											0.402	***		
log(EP_0): Dummy(MAC)											0.327	**		
log(EP_0): Dummy(TRE)											-0.037			
log(EP_0): Dummy(NSI)											0.323	**		
log(EP_0): Dummy(BEL)													0.098	
log(EP_0): Dummy(CZE)													0.073	
log(EP_0): Dummy(DNK)													0.090	
log(EP_0): Dummy(FIN)													0.169	**
log(EP_0): Dummy(FRA)													0.113	
log(EP_0): Dummy(GER)													0.089	
log(EP_0): Dummy(HUN)													0.185	**
log(EP_0): Dummy(ITA)													0.091	
log(EP_0): Dummy(NLD)													0.039	
log(EP_0): Dummy(POL)													0.013	
log(EP_0): Dummy(PRT)													0.037	
log(EP_0): Dummy(SPA)													0.100	
log(EP_0): Dummy(SVK)													0.207	***
log(EP_0): Dummy(SWE)													0.190	**
log(EP_0): Dummy(UK)													0.141	*
ρ	0.364	***	0.353	***	0.354	***	0.307	***	0.354	***	0.216	**	0.418	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.3: Estimation Results of Fixed Effect Spatial Error Models under "Pooled Effects"

	Case 1		Case 2		Case 3		Case 4		Case 5		Case 6		Case 7	
Intercept	-3.719	***	-3.524	***	-1.622		-3.690	***	-4.007	***	-1.069		0.011	
log(EP_0)	-0.030	**	-0.027	*	-0.029	**	-0.038	*	-0.012		-0.388	***	-0.183	***
EPRIC	0.123	***	0.042		0.048		0.128	***	0.135	**	0.182	***	0.139	***
log(HDDs + CDDs)	0.322	***	0.413	***	0.206	***	0.325	***	0.3209	***	0.271	***		
K_SHARE	-1.296	***	-1.543	***	-1.542	***	-1.282	***	-1.379	***	-1.307	***	-1.738	***
LEAD			1.853	***										
NE					-0.215	***								
log(EP_0): Dummy(T2)							0.018							
log(EP_0): Dummy(H_EPRIC)									-0.083	***				
log(EP_0): Dummy(TEX)											0.083			
log(EP_0): Dummy(WOD)											0.306	**		
log(EP_0): Dummy(PAP)											0.217			
log(EP_0): Dummy(CHE)											0.377	**		
log(EP_0): Dummy(NMM)											0.070			
log(EP_0): Dummy(BMI)											0.399	***		
log(EP_0): Dummy(MAC)											0.323	**		
log(EP_0): Dummy(TRE)											-0.042			
log(EP_0): Dummy(NSI)											0.323	**		
log(EP_0): Dummy(BEL)													0.077	
log(EP_0): Dummy(CZE)													0.105	
log(EP_0): Dummy(DNK)													0.117	
log(EP_0): Dummy(FIN)													0.181	**
log(EP_0): Dummy(FRA)													0.079	
log(EP_0): Dummy(GER)													0.071	
log(EP_0): Dummy(HUN)													0.213	***
log(EP_0): Dummy(ITA)													0.077	
log(EP_0): Dummy(NLD)													0.015	
log(EP_0): Dummy(POL)													0.025	
log(EP_0): Dummy(PRT)													0.032	
log(EP_0): Dummy(SPA)													0.061	
log(EP_0): Dummy(SVK)													0.234	***
log(EP_0): Dummy(SWE)													0.222	***
log(EP_0): Dummy(UK)													0.155	*
λ	0.358	***	0.347	***	0.356	***	0.307	***	0.360	***	0.168	*	0.484	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

From table 2.2 and table 2.3, it can be seen that the almost all estimates of β (namely, coefficient of *log(EP_0)*) are negative and statistically significant at the 1% level. Concerning the additional explanatory variables, it is found that energy price has the expected (positive) sign and shown statistically significance in most cases, indicating that high energy prices have a positive impact on energy productivity growth. The weather condition variable (logarithm of the sum of *HDDs* and *CDDs*) also has positive and statistically significant impact on energy productivity growth. The implication is that for countries facing high energy prices or severe weather conditions that demand energy for cooling or heating would generate strong incentives to improve their energy productivity levels. In terms of environmental policy stringency variable "*LEAD*", in case 2, the estimated coefficient turns to be positive and statistically significant. It implies that countries with less stringent environmental policy (namely, with high lead content in gasoline) tend to witness faster energy productivity growth. Such a finding makes me wonder whether the "*LEAD*" variable directly correlates with the development stage of the EU countries. Countries with low lead content in gasoline or high environmental stringency are always the countries at their relatively developed stage. Due to this concern, "*LEAD*" variable is replaced by the dummy variable "*NE*" in case 3 and the estimate associated with "*NE*" turns out to be negative and significant. In other words, non-eastern EU countries are slow in increasing energy productivity when compared to less developed eastern EU countries. It reaffirms my speculation on technology catching up that the rate of energy productivity improvement is contingent upon the development stage of the country. This is also true when it comes to interpreting the estimated coefficients for variable "*K_SHARE*". High capital share for the production implies a high level of technology advancement and capital can also substitute the inputs of raw energy materials, and thus for these countries with high capital share they witness a relatively slower energy productivity growth rate.

Estimated coefficient for the time period dummy variable (T2: 2000-2005) in case 4 is positive, showing that at the second stage of the study period the energy productivity growth is faster. Nevertheless, neither of time-associated dummies for the two models (spatial lag and spatial error) turns out to be statistically significant. In case 5, estimated β coefficients for countries with higher energy price (using dummy variable *H_EPRIC*) for both models are negative and significant, while the default β coefficients (for the entire sample) are not significant. This indicates that convergences in energy productivity are mostly found among countries with relatively high energy prices. It again proves that rising energy price has a positive pressure on energy efficiency catching up. Moreover, in order to understand the differences of energy productivity growth rates for each manufacturing sector, the estimated coefficients associated with the sector dummies can provide some insights. In case 6, five detailed manufacturing sectors including BMI (*Basic Metals*), CHE (*Chemistry*), MAC (*Machinery*), NSI (*Non-Specified Industry*), and WOD (*Wood and of Wood and Cork*) show statistically significant effect of energy productivity convergence. This makes sense since most of these are energy-intensive sectors and they face strong needs to improve energy productivity. In similar fashion, in case 7, country-fixed effects of energy productivity convergence are found to statistically significant in Finland (FIN), Hungary (HUN), Slovakia (SVK), Sweden (SWE), and United Kingdom (UK).

Two different specifications (spatial lag and spatial error) of the spatial panel models produce very similar estimation results under the "pooled effects". The estimated spatial autoregressive and autocorrelation coefficients (ρ and λ) are both positive and statistically significant. It proves that the energy productivity growth of one country is positively related to the growth of its trading partner countries. The trade-facilitated technology spillovers mostly occur among countries that share strong intra-industry trade flow partnership.

Table 2.4 and table 2.5 present the estimation results when "cross-section specific effects" and "both spatial and time period fixed effects" are considered in the fixed effects spatial lag and error models. It is worth to notice that coefficients of time-invariant variables cannot be estimated under such two scenarios.

Thus, variables such as "*log(HDDs + CDDs)*", "*LEAD*", and all the dummy variables have to be dropped in the estimated equation. The estimates of β are negative and statistically significant, indicating a converging process of energy productivity level across EU countries. However, these estimated βs are in a much higher magnitude than the estimates in table 2.2 and 2.3. Mostly because there are still varying effects associated with country-specific and industry-specific factors that are not captured in fixed effects models under "cross-sectional specific effects" and "both spatial and time fixed effects". The estimated coefficients for "*K_SHARE*" are negative and significant, and similar interpretations can be made as discussed above in the "pooled effects" estimates for "*K_SHARE*". The estimated spatial autoregressive coefficients for ρ are positive but not statistically significant, but the autocorrelation coefficients for λ are all positive and statistically significant. It shows that the trade-induced correlations (or spillovers) are mostly found in the error terms of the panel convergence models, indicating nuisance error dependence in the energy productivity convergence process.

Table 2.4: Estimation Results of Fixed Effect Spatial Lag Models under "Cross-Sectional Specific Effects" and "Both Spatial and Time Fixed Effects"

	Cross Sectional Specific Effects		Both Spatial and Time Period Fixed Effects	
log(EP_0)	-0.738	***	-0.748	***
EPRIC	0.035		0.015	
K_SHARE	-1.592	***	-1.498	**
ρ	0.101		0.125	

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.5: Estimation Results of Fixed Effect Spatial Error Models under "Cross-Sectional Specific Effects" and "Both Spatial and Time Fixed Effects"

	Cross Sectional Specific Effects		Both Spatial and Time Period Fixed Effects	
log(EP_0)	-0.756	***	-0.764	***
EPRIC	-0.003		-0.016	
K_SHARE	-1.488	**	-1.436	**
λ	0.233	***	0.239	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.6 shows estimation results of the spatial error model when spatial correlation and random effects are considered. Random effects model avoids the loss of degrees of freedom incurred in the fixed effects model and the problem that the coefficients of time-invariant variables cannot be estimated. The findings are very similar to the fixed effect spatial error model under "pooled effects". Almost all the conditional variables are with the expected sign and the estimates of βs are negative and statistically significant, indicating a conditional convergence process. In case 4, it again shows that countries with higher energy prices are witnessing strong energy productivity converging process, and the estimated rates (-0.098) are generally higher than the default overall βs estimates in case 1, 2, and 3 (-0.041, -0.036, and -0.042). The associated spatial autocorrelation coefficients λ are positive and statistically significant, showing the energy productivity growth of one country is positively related to its trading partners and technology spillovers are embedded in the trade flows.

Table 2.6: Estimation of Error Models under "Spatial Correlation and Random Effects"

	Case 1		Case 2		Case 3		Case 4	
Intercept	-3.685	***	-1.613		-3.780	***	-3.848	***
log(EP0)	-0.041	**	-0.036	**	-0.042	**	-0.02	
EPRIC	0.125	**	0.051		0.133	***	0.1229	**
log(HDDs + CDDs)	0.330	***	0.210	***	0.336	***	0.3322	***
K_SHARE	-1.549	***	-1.711	***	-1.557	***	-1.654	***
NE			-0.216	***				
log(EP0): Dummy(T2)					-0.127			
log(EP0): Dummy(H_EPRIC)							-0.098	***
λ	0.393	***	0.386	***	0.316	***	0.398	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

2.4.4 Estimation Results with Distance Weight Matrix

For identification purposes, the spatial weight matrix needs to be defined exogenously (Manski, 1993). The choice of an appropriate spatial weight matrix is an important methodological problem in the quantitative spatial dependence literature and is subject to some arbitrariness. Such arbitrariness could present a serious problem to inference model results, since estimation could depend critically on the choice of spatial weight matrix (Anselin, 2002; Fingleton, 2003). Hence, it becomes worthwhile to consider the performance of some alternative specifications of the weight matrix. In this chapter, the weighting scheme could be based on either spatial distance or economic distance. The trade flow weight matrix discussed above is a good example of economic distance. Geographical distances are also often used to approximate the strength of social relationships. Therefore, I re-estimate the energy convergence panel model using the traditional distance weight matrix to check the sensitivity of weight matrix specification on the results and corresponding interpretations.

Table 2.7: Estimation Results of Fixed Effect Spatial Lag Models under "Pooled Effects" (Using Distance Weight Matrix)

	Case 1		Case 2		Case 3		Case 4		Case 5		Case 6		Case 7	
Intercept	-3.899	***	-3.910	***	-1.920	**	-3.995	***	-4.070	***	-1.096		0.082	
log(EP0)	-0.030	**	-0.026	*	-0.028	*	-0.029		-0.011		-0.392	***	-0.164	**
EPRIC	0.124	***	0.053		0.057		0.135	***	0.128	***	0.188	***	0.118	***
log(HDDs + CDDs)	0.339	***	0.437	***	0.223	***	0.339	***	0.334	***	0.271	***		
K_SHARE	-1.163	***	-1.379	***	-1.379	***	-1.168	***	-1.226	***	-1.359	***	-1.529	***
LEAD			1.881	***										
NE					-0.216	***								
log(EP0): Dummy(T2)							-0.112							
log(EP0): Dummy(H_EPRIC)									-0.083	***				
log(EP0): Dummy(TEX)											0.089			
log(EP0): Dummy(WOD)											0.299	**		
log(EP0): Dummy(PAP)											0.218			
log(EP0): Dummy(CHE)											0.399	**		
log(EP0): Dummy(NMM)											0.057			
log(EP0): Dummy(BMI)											0.402	***		
log(EP0): Dummy(MAC)											0.319	**		
log(EP0): Dummy(TRE)											-0.045			
log(EP0): Dummy(NSI)											0.322	**		
log(EP0): Dummy(BEL)													0.104	
log(EP0): Dummy(CZE)													0.074	
log(EP0): Dummy(DNK)													0.086	
log(EP0): Dummy(FIN)													0.153	*
log(EP0): Dummy(FRA)													0.115	
log(EP0): Dummy(GER)													0.089	
log(EP0): Dummy(HUN)													0.188	**
log(EP0): Dummy(ITA)													0.093	
log(EP0): Dummy(NLD)													0.041	
log(EP0): Dummy(POL)													0.004	
log(EP0): Dummy(PRT)													0.045	
log(EP0): Dummy(SPA)													0.106	
log(EP0): Dummy(SVK)													0.208	***
log(EP0): Dummy(SWE)													0.195	**
log(EP0): Dummy(UK)													0.146	*
ρ	0.055		0.091		0.104		0.009		0.054		-0.026		0.197	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.8: Estimation Results of Fixed Effect Spatial Error Models under "Pooled Effects" (Using Distance Weight Matrix)

	Case 1		Case 2		Case 3		Case 4		Case 5		Case 6		Case 7	
Intercept	-4.096	***	-3.848	***	-1.803	*	-4.088	***	-4.349	***	-1.048		0.090	
log(EP0)	-0.031	**	-0.026	*	-0.028	*	-0.030		-0.012		-0.395	***	-0.188	***
EPRIC	0.127	***	0.045		0.051		0.135	***	0.137	***	0.190	***	0.135	***
log(HDDs + CDDs)	0.360	***	0.444	***	0.222	***	0.351	***	0.356	***	0.265	***	-1.611	***
K_SHARE	-1.194	***	-1.456	***	-1.473	***	-1.184	***	-1.256	***	-1.380	***		
LEAD			1.933	***										
NE					-0.221	***								
log(EP0): Dummy(T2)							-0.119							
log(EP0): Dummy(H_EPRIC)									-0.083	***				
log(EP0): Dummy(TEX)											0.088			
log(EP0): Dummy(WOD)											0.298	**		
log(EP0): Dummy(PAP)											0.219			
log(EP0): Dummy(CHE)											0.405	**		
log(EP0): Dummy(NMM)											0.054			
log(EP0): Dummy(BMI)											0.404	***		
log(EP0): Dummy(MAC)											0.329	**		
log(EP0): Dummy(TRE)											-0.047			
log(EP0): Dummy(NSI)											0.325	**		
log(EP0): Dummy(BEL)													0.1167	
log(EP0): Dummy(CZE)													0.1052	
log(EP0): Dummy(DNK)													0.1071	
log(EP0): Dummy(FIN)													0.1765	**
log(EP0): Dummy(FRA)													0.1166	
log(EP0): Dummy(GER)													0.1003	
log(EP0): Dummy(HUN)													0.212	***
log(EP0): Dummy(ITA)													0.1049	
log(EP0): Dummy(NLD)													0.0492	
log(EP0): Dummy(POL)													0.0319	
log(EP0): Dummy(PRT)													0.0574	
log(EP0): Dummy(SPA)													0.1234	
log(EP0): Dummy(SVK)													0.2424	***
log(EP0): Dummy(SWE)													0.2145	***
log(EP0): Dummy(UK)													0.1702	**
λ	0.102		0.152	**	0.164	**	0.053		0.109		-0.056		0.240	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.9: Estimation Results of Fixed Effect Spatial Lag Models under "Cross-Sectional Specific Effects" and "Both Spatial and Time Fixed Effects" (Using Distance Weight Matrix)

	Cross Sectional Specific Effects		Both Spatial and Time Period Fixed Effects	
$log(EP_0)$	-0.756	***	-0.755	***
EPRIC	0.030		0.031	
K_SHARE	-1.656	***	-1.661	**
ρ	-0.025		-0.026	

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.10: Estimation Results of Fixed Effect Spatial Error Models under "Cross-Sectional Specific Effects" and "Both Spatial and Time Fixed Effects" (Using Distance Weight Matrix)

	Cross Sectional Specific Effects		Both Spatial and Time Period Fixed Effects	
$log(EP_0)$	-0.760	***	-0.760	***
EPRIC	0.087		0.088	
K_SHARE	-2.175	***	-2.175	**
λ	-0.277	***	-0.276	***

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

Table 2.11: Estimation of Error Models under "Spatial Correlation and Random Effects" (Using Distance Weight Matrix)

	Case 1		Case 2		Case 3		Case 4	
Intercept	-3.936	***	-1.740	*	-4.074	***	-4.086	***
log(EP0)	-0.042	**	-0.034	**	-0.035	*	-0.019	
EPRIC	0.125	***	0.052		0.139	***	0.124	**
log(HDDs + CDDs)	0.360	***	0.221	***	0.354	***	0.357	***
K_SHARE	-1.474	***	-1.627	***	-1.531	***	-1.565	***
LEAD								
NE			-0.221	***				
log(EP0): Dummy(T2)					0.019			
log(EP0): Dummy(H_EPRIC)							-0.099	***
λ	0.145	*	0.183	**	0.059		0.148	*

*(Asterisks denote levels of Significance: *** at 1%, ** at 5%; * at 10 %.)*

From table 2.7 to table 2.11, it is easy to notice that the estimated coefficients associated with the conditional variables are very similar to my previous results using the trade flow weight matrix. However, significant differences are found in the estimates of the spatial autoregressive and autocorrelation coefficients (ρ and λ). In the case of using trade flow weight matrix, the estimates of ρ and λ are almost all positive and significant, indicating technology spillover effects among trading partner countries. However, in the case of using distance weight matrix, the estimates of ρ and λ are with mixed signs (either positive or negative), and most of the estimates are not statistically significant. This means physical distance might not serve as a good instrument to capture the technological communication among EU countries, in that geographically close countries might not witness strong technology diffusion. Instead, trade flow among EU countries could account for energy productivity spillover effects more.

2.5 Conclusions and Discussions

This chapter has offered some empirical spatial panel analyses of energy productivity convergence by specifically testing the technology spillover effects associated with trade flow among 16 EU countries, covering the period 1995-2005. Under the panel-data approach, the estimated coefficients of β are almost all negative and statistically significant. It leads to the conclusion that EU countries with low starting energy productivity levels witness relatively faster growth of energy productivity, and it further proves that relatively backward EU countries tend to catch up to more advanced economies, particularly the energy productivity levels.

The estimation results also provide evidence that energy productivity convergence is conditional upon the cross-country differences in steady-state characteristics. More specifically, it is found that high energy prices and severe weather conditions tend to stimulate energy productivity growth. A negative relationship is found between the capital share of the industrial production process and the energy productivity growth. Non-eastern EU countries show statistically significant lower rates of energy productivity growth. And energy use intensive manufacturing sectors (such as Basic Metals, Chemicals, Machinery, Wood and of Wood and Cork) tend to show statistically significant rates of energy productivity converging effects. However, the results fail to show statistically significant difference in the energy productivity growth rates between 1995-2000 and 2000-2005.

The application of exploratory spatial panel data analysis methods reveals strong evidence of trade-flow-induced spatial correlation in the energy productivity growth process. Most of the estimated autocorrelation coefficients (ρ and λ) tend to be positive and statistically significant. This further means that while EU countries may be converging in country- and industry-specific energy productivity terms, they do not do so independently but rather to display movements similar to their trading partner countries. It also shows that technological externalities associated with international intra-industry trade flows should receive greater attention in the global warming and energy economics fields. High energy price could be a significant factor to induce technology innovation and subsequently promote spillover effects.

Taking advantage of the presence of spatial correlation in energy productivity improvement process, policy makers could simulate the effects of random shocks (e.g. innovation or adopting new efficient technologies) to individual manufacturing sectors to not only move that country away from its current energy productivity level, but also to propagate the effects throughout the spatial correlation system to other EU countries. It helps project energy productivity at sectoral level across all concerned EU countries, and will be especially meaningful in designing EU-wide energy policies, such as the well-known EU ETS. Emissions projections have always been of vital importance to estimate the costs of emission trading programs, but unfortunately there were very few sector-specific emission prediction models (Ellerman and Buchner, 2007). The projection of energy productivity levels across EU countries at the detailed sector level, in combination with other factors such as expected rates of growth in economic activity, energy types, and effects of regulatory provisions can address the issue of predicting an inherently uncertain future of carbon emission level. Effective forecast of emission becomes a basis for the allocation of European Union Allowances (EUAs) in the future.

The major limitations of this chapter arise from the following sources. First, the estimation results heavily rely on the recently developed spatial panel regression package in R. There are important tests to be further developed, such as Hausman test to select the preferred model between random effects and fixed effects in a panel setting, and the estimation approach for the spatial lag random effects model under panel structure. I am very confined to the available functions in this package to generate estimation results. Second, there still lacks efficient statistical tests to estimate the choice of a spatial weight matrix amongst a predetermined set of alternatives (in this chapter I tested two - trade flow based weight matrix and distance based matrix - and there are many other options). These limitations will be left for future work to improve.

Chapter 3

GROUNDWATER RESOURCE PLANNING TO PRESERVE STREAMFLOW - WHERE ENVIRONMENTAL AMENITY MEETS ECONOMIC WELFARE LOSS

3.1 Introduction

Groundwater and surface water are not isolated, but inter-connected through an identified aquifer system in many regions. Excessive groundwater pumping, especially during drought periods, may cause the natural flow in the streams to fall below a level necessary for instream uses. As a result, streamflow may dry up and jeopardize a healthy ecologic and aquatic community, leading to reduced environment amenities. Federal and State legislation has stated its intention to promote sustainable economic growth while protecting the natural environment by minimizing the occurrence of low streamflows (Viscusi, 1995).

Conjunctive groundwater/surface water models have been proposed to manage the water resources, and some of them have paid specific attention to stream depletion (Theis, 1941; Glover and Balmer, 1954; Hantush, 1965; Hunt, 1999). A large body of work adopted simulation and optimization methods for some particular groundwater systems to provide insights for groundwater management (Bredehoeft and Young, 1970; Gorelick, 1983). Maddock (1974) developed operating rules that relate streamflow interaction with well pumping and with the statistics of the demands, streamflow, pumping and drawdowns. Mueller (1993) established a model to link groundwater withdrawals with surface streamflow and optimally implemented the permit conditions so that the depletion of streamflow stays below a certain standard. Some recent work has been devoted to modifying the hydrologic assumptions or aquifer conditions in previous stream-aquifer-system studies and re-estimating depletion rates/volume (Butler *et al*., 2007; Yeh *et al*., 2008; Ward and Callander, 2010). These models help ascertain the spatial and temporal interactions of the stream-well-aquifer system.

In addition to studies on physical hydrological mechanisms, the economic impacts associated with groundwater supply have also been incorporated into the decision framework of groundwater management. Young and Bredehoeft (1972) developed a simulation model composed of a hydrologic model and an economic model. The latter one represented the responses of irrigation water users to variations in water supply and cost. Feinerman and Knapp (1983) investigated the benefits of groundwater management and their related welfare effects on groundwater users. They showed that socio-economic impacts caused by groundwater use regulations may also vary over space.

Effective and holistic management of groundwater resources would ideally incorporate concerns of preserving environmental amenities while minimizing the negative socio-economic impacts. In this chapter, tradeoffs between restoring streamflow and economic welfare loss at a small spatial scale are analyzed. It proposes an integrated hydrologic-economic framework to study groundwater resources at county and municipality levels. The tradeoffs are based on the locations of the wells and streams, physical hydrologic factors, and water rates. The rest of the chapter is organized as follows. In section 3.2, I begin by introducing the groundwater use facts in the study site. Following the background, components of the optimization model, including decision variables, objectives and constraints, are elaborated. It also contains a brief description of the genetic algorithm I intend to use in this chapter. Section 3.3 explains the selection of benchmark year for this analysis and the available sources of data. Section 3.4 presents the Pareto frontiers under both county and municipality levels. Section 3.5 discusses the modeling results and the sensitivities of return flow coefficients and pumping boundary constraints I impose in the modeling process. The last section 3.6 summarizes and identifies some possible directions for future research.

3.2 Methodology

3.2.1 Study Site

The study site, McHenry County (see figure 3.1), is located approximately 55 kilometers northwest of downtown Chicago. In the past several decades, McHenry County has undergone a fast urbanization and suburbanization process. From 1969 to 2008, the population almost tripled from 0.1 million to 0.3 million (Data Source: Regional Economic Accounts, Bureau of Economic Analysis). In McHenry County, groundwater functions as an especially key resource to support regional economic development as underlying aquifers provide almost all of the county’s water supply. In response to the rocketing population, total demand for groundwater has also increased dramatically. The groundwater withdrawals in McHenry County had risen from 1.2 billion cubic feet in 1979 to 2.3 billion cubic feet in 2008 (Data Source: Illinois State Water Survey). Immoderate withdrawal of groundwater can speed up the process of stream depletion in this area. From a land use perspective, the entire land cover area of McHenry County is 610.7 square miles, and in 2001 the highest portion (57%) of land was for agricultural use. Between 2001 and 2005, comparing to other categories of land use types, agricultural land cover suffered the most severe decrease from 348.12 square miles to 321.64 square miles (Data Source: Land Use Inventory, Chicago Metropolitan Agency for Planning). As agricultural land and wetlands are being converted to impervious surfaces, groundwater recharge rates may be lowered, which further exacerbates the streamflow depletion problem. According to the Illinois State Water Survey Groundwater Database, a significant portion (85%) of water supplied by the groundwater resources in McHenry County goes to public water supply, mostly residential water use. Therefore, the focus of this chapter is on groundwater supply for domestic use.

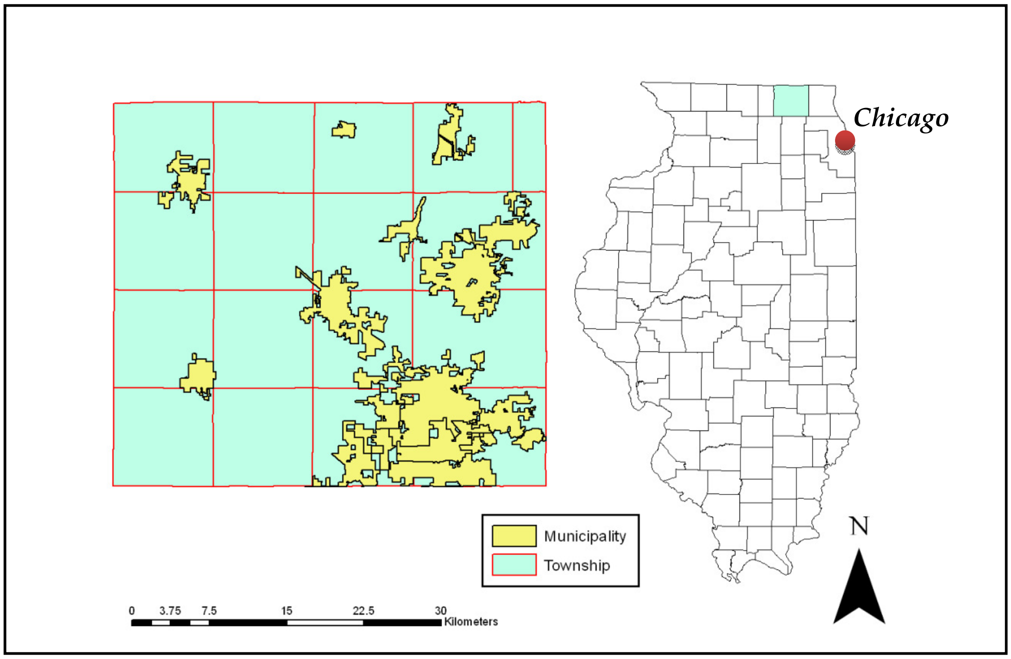

Figure 3.1: Location of McHenry County

3.2.2 Streamflow Depletion Due to Groundwater Pumping

With the initiation of groundwater withdrawals, groundwater levels in the aquifer will start decreasing and some of the groundwater that flows to the stream in the absence of withdrawals is now captured by the wells. If the pumping rate is high enough, instream flow could be drawn into the aquifer ("infiltration") and possibly captured by the well. Reduction in streamflow caused by groundwater withdrawals are considered as streamflow depletion. Efforts on calculation of stream depletion have been made by numerous scholars. Theis (1941) is considered as the first scholar to assess the impact of groundwater pumping on an associated stream. Glover and Balmer (1954) and later Jenkins (1968) developed a more generalized analytical approach that substituted the earlier integral formulation of Theis with the complementary error function, which serves as a more useful means to describe the complete aquifer – well relationship. For a single well, the analytical solution offered by Glover and Balmer (1954) to calibrate the stream depletion due to pumping in an aquifer is:

$$D(a, S, T) = Q * erfc\left(\sqrt{\frac{a^2 S}{4Tt}}\right) \tag{3-1}$$

where *erfc* is the complementary error function; D [L^3/T], flow depletion taken from the stream; a [L], distance between the well and the stream; S [Dimensionless], storativity of a confined aquifer or specific yield of an unconfined aquifer; T [L^2/T], transmissivity of the aquifer (assuming homogenous between the well and the stream); Q [L^3/T], daily pumpage; t [T], time elapsed since the start of pumping, and in this chapter I use 365 days. The total daily depletion of stream (*Tot_D*) can be obtained through:

$$Tot\_D = \sum_n D_n(a, S, T) = \sum_n \left[Q_n * erfc\left(\sqrt{\frac{a_n^2 S_n}{4T_n t}}\right)\right] \tag{3-2}$$

where n is the well index.

3.2.3 Return Flow to the Streams

Irrevocable water loss occurs when it is consumed by people or animals, transpired by plants, evaporated, or incorporated into plants or products (Solley *et al.*, 1998). Consumptive water use may constitute a small portion of the total water withdrawal, and non-consumed water is returned to the hydrologic system as return flows. Therefore, calibration of the amount of return flow back to the stream is an important component to obtain the final balance of streamflow quantity. Usually, public-disposed wastewater is collected through sewers or wastewater-collection systems, treated at a wastewater treatment plant (WWTP) and discharged to receiving water such as a river, estuary, or aquifer. A WWTP serves one or more municipalities and is permitted under the U.S. Environmental Protection Agency (EPA) National Pollutant Discharge Elimination System (NPDES).

Return flows to the streams in McHenry County are mainly estimated by information on discharges obtained from these WWTPs. Another source of return flow is from wastewater disposed by individual users. However, due to the difficulty in obtaining such information, return flows from self-disposed wastewater through on-site septic systems are not considered in this chapter. Also, with no detailed information on where WWTPs receive their wastewater input, it is reasonable to assume that the processing efficiency of wastewater by WWTPs is the same across all 15 municipalities in McHenry County. The return flow coefficient is then calculated by the ratio of effluent discharged from all WWTPs to the total groundwater withdrawal on a daily basis (see equation 3-3). The coefficient is used as a proxy for the percentage of the depleted streamflow that will be eventually returned to the streams in a typical water cycle.

$$Coeff\_Return = \frac{Discharge\_WWTP}{\sum_n Q_n} \times 100\% \quad (3\text{-}3)$$

3.2.4 Economic Welfare Loss

According to equation (3-2), when the physical hydrogeological characteristics are fixed, reducing pumpage Q_n is one way to reduce stream depletion. However, restricting pumping (in other words, reducing water supply) leads to negative economic impacts. As was mentioned earlier, a significant portion of groundwater supply goes to residential use. Unlike industrial production processes, the residential sector does not yield a directly measurable monetary cost when facing water restrictions. Usually, losses are measured by consumers' willingness to pay to avoid water service interruptions, defined as the amount of money that residential customers would pay in order to avoid a break in water service of some duration (Brozovic *et al.*, 2007; Jenkins *et al.*, 2003). Scholars have explored several different approaches to estimate the economic losses associated with water supply shortage scenarios. A detailed analysis on welfare loss under water rationing scenarios can be found in Wan *et al*. (2011). In this chapter, I adopt the approach of Jenkins *et al*. (2003) by formulating water demand equations to estimate the willingness to pay from residential end users' side to avoid water shortage.

The elasticity denotes the change of marginal quantity of water consumed in case of a marginal water price alteration. The price elasticity of demand (η) for water in the residential sector can be defined by:

$$\eta = (dD/D)/(dP/P) \approx (\triangle\, D/D)/(\triangle\, P/P) \quad (3\text{-}4)$$

where P is the unit price of residential water (dollars per thousand cubic feet) and D is residential water demand (thousand cubic feet per day) at that price. The demand for water is said to be "inelastic" or "elastic" with regard to price depending on whether the value of the price-elasticity parameter lies between 0 and -1 or below -1. When the price of water with an inelastic demand rises, more is spent on water than previously because the decrease in consumption is insufficient to offset the increase in water price. On the other hand, when the price of a commodity with an elastic demand rises, less is spent on it than previously because the decrease in consumption more than offsets the increase in water price (Lesser, 1954).

When supply is restricted, users have to bid higher prices to gain access to limited water resources. Values on water are calibrated under the sense of how much the residential users (households) are willing to pay for each additional unit of water along their demand curves. When water supply is unrestricted, following the demand curve (see figure 3.2), residential users are willing to pay price P_1 by consuming the ideal amount of water (D_{max}). Under the intension to alleviate the pressure on the stream depletion, the amount of water to be supplied (D_{actual}) is restricted to be less than the optimal unconstrained quantity ($D_{actual} < D_{max}$). Thus, the economic loss represents the economic value or benefits that users would gain from additional water if deliveries were increased to the maximum quantity demanded.

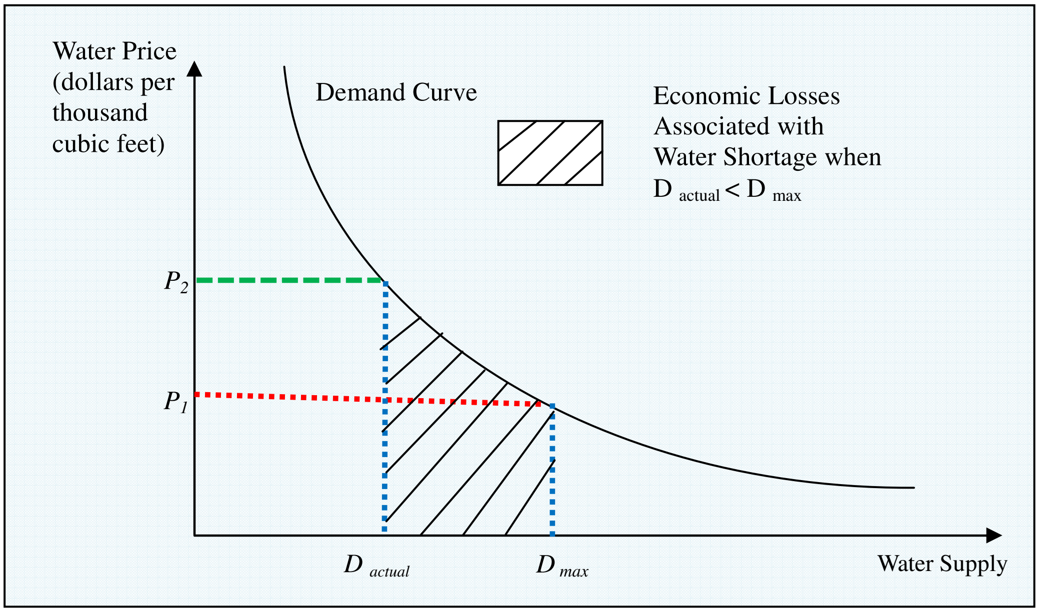

Figure 3.2: Demand Curve for Water Consumption in Residential Sector

If I assume a constant price elasticity of demand (η), then equation (3-4) can be integrated to give the residential water demand function as:

$$P_1 = \exp\left[(lnD_{max}/\eta) + c\right] \tag{3-5}$$

in which c is an integration constant and can be estimated by:

$$c = \ln(P_1) - [\ln(D_{max})/\eta] \tag{3-6}$$

Usually, surveys are sent to households or individuals to ask for their willingness to pay for water services or their likely reaction to changed water prices. Elasticity of demand for water can then be calculated from the survey data. If elasticity estimates are available, then the demand function can be estimated from the available price and quantity of water use. In other words, P, D, and η are the three parameters needed to establish/locate the residential demand curve. Economic loss can be calculated by integrating the demand curve from maximum residential demand leftward to the actual delivery. In figure 3.2, the shaded area indicates the welfare loss due to water supply constraint. After integration, the economic loss can be expressed as:

$$Econ\_Loss = [\exp(c)/(1 + 1/\eta)] \times \left[D_{max}^{(1+1/\eta)} - D_{actual}^{(1+1/\eta)}\right] \tag{3-7}$$

3.2.5 *Multi-Objective Decision Making Model*

Environmental management problems always have multiple objectives and are spatially complex and potentially conflicting. In this study, the physical response of streamflow to groundwater withdrawals determines the planning strategies that may be used to balance residential water supply and aquatic-habitat protection goals. More specifically, the multi-objective groundwater resource management model consists of a set of decision variables, two objective functions, and a set of constraints. The decision variables are daily pumping rate at each individual well (Q_n) that supplies water for residential use. Let's assume U_n as the current pumping rate without any restriction for the well n in McHenry County, and the constrained desirable pumping rate is R_n (both in the unit of thousand cubic feet per day). The goal is to minimize stream depletion (in other words, restore streamflow) and also to minimize the economic welfare loss produced by constraining well pumpage. Equation (3-8) estimates the amount of daily streamflow depletion/loss in McHenry County by considering the water balances between outflows (depletion through groundwater pumping) and inflows (wastewater discharge into streams). Equation (3-9) shows the total economic welfare loss associated with restrictions on pumping. Thus, the objectives of this decision-making process are to understand the tradeoffs between streamflow depletion and total economic welfare loss.

$$Stream\_Loss = (1 - Coeff\_Return) * \sum_n \left[R_n * erfc\left(\sqrt{\frac{a_n^2 S_n}{4T_n t}}\right)\right] \tag{3-8}$$

$$TOT\_ECON\_LOSS = \sum_{i=1}^{n} [\exp(c_n)/(1 + 1/\eta)] \times \left[U_n^{(1+1/\eta)} - R_n^{(1+1/\eta)}\right] \tag{3-9}$$

In addition to equations (3-8) and (3-9), two constraints are imposed in the multi-objective optimization process. Equation (3-10) specifies the lower bound (*LB*) of the restricted pumping rate at each well. LB_n is a factor assigned to each well based upon its magnitude of impact toward stream depletion. From equation (3-1), I know that the contribution of each well's pumping toward stream depletion depends on the pumping rate and a well-specific depletion factor $erfc\left(\sqrt{\frac{a_n^2 S_n}{4T_n t}}\right)$. In this study, the depletion factors for all wells are calculated and ranked from highest to lowest. The wells with the 25% highest depletion factors are assigned a zero value for the lower bound variable *LB*. In the model, these wells can be shut down as R_n equals zero if necessary in the worst case. For the second (25%~50%), third (50%~75%), and fourth quarter (75~100%) of the depletion factor rank (high to low), the *LB*s are assigned values of 40%, 50%, and 60%. The 40% lower bound was chosen on the fact that 40 gallons is considered as a necessity for a socially and economically developed community. The regular per capita domestic water consumption in McHenry County is close to 100 gallons per day (Wan *et al.*, 2011). The restriction becomes loose when the depletion factor becomes low. For the wells ranked in the lowest quarter of the depletion factor range, the constrained pumping must always stay above 60% of their previous unconstrained pumping rates. Equation (3-11) ensures the upper bound that the regulated pumping should stay below the previous unconstrained level of pumping. For practical implementation, the classifications and values of LBs can be easily adjusted accordingly.

$$R_n \geq U_n * LB_n \tag{3-10}$$

$$R_n \leq U_n \tag{3-11}$$

3.2.6 *Implementation of NSGA-II Method for Multi-Objective Mathematical Programming*

Unlike the single-objective optimization problems whose solutions are usually straightforward, a multi-objective mathematical problem has no single optimal solution that can simultaneously optimize all the objective functions. This is especially true for environmental management decision-making processes, where multiple objectives may conflict. The challenge is to evaluate the performance of the action relative to the multiple objectives. Decision makers might have to explore a full set of solutions instead of a single answer.

Frequently used is an a posteriori method that approximates the non-dominated Pareto frontier without using weightings and allows for visualization of the tradeoffs among objectives. A solution is considered as non-dominated if there exists no other feasible solution that will give an improvement in one objective without a subsequent degradation in at least one other objective (Cohon, 1978). The decision makers can then express their preferences by analyzing objective tradeoffs and select a point on the Pareto frontier as the most preferred one for environmental planning and management purpose. However, given the size of decision variables and objectives, consideration of the entire feasible space to determine the non-dominated Pareto frontier can be computationally burdensome. The genetic algorithm (GA) has emerged as a robust global optimization algorithm and has been applied to many water quality management problems (Kalwij, 2008). GA searches complex problem spaces by using a process that is analogous to Darwinian natural selection (Srinivas and Deb, 1994). GA is particularly suitable for multi-objective optimization problems, since it deals simultaneously with a set of possible solutions and produces Pareto optimal solutions in one single run of the algorithm.

In the context of this chapter, a fast and elitist multi-objective genetic algorithm is selected to integrate the search algorithm with the calculation of objectives to locate the non-dominated Pareto frontier. The optimization framework is based on a multi-objective optimization program, Non-dominated Sorting Genetic Algorithm (NSGA-II), developed by Deb *et al*. (2002). The NSGA-II algorithm is composed of five operators: initialization, fast non-dominated sorting, crossover, mutation, and the elitist crowded comparison operator. Different from other evolutionary algorithms, NSGA-II uses the non-dominated sorting and ranking selection with the crowded comparison operator (Deb *et al.*, 2002). A real-coded NSGAII is employed as the optimization mechanism to search the solution space for the fittest individuals based on two objective functions while maintaining optimal individuals non-dominated by any other solutions. A source code of NSGA-II written by Aravind Seshdri (Matlab Exchange, 2007) was modified to incorporate the simulation model and objective functions in this chapter.

3.3 Data

3.3.1 Simulation Benchmark Year

Maintaining an appropriate level of instream flow is needed to protect aquatic and riparian ecosystems. However, minimal ecological requirements for instream flow are not well defined. In most cases, criteria of instream flow are developed from historical streamflow statistics or regional average streamflow statistics. Due to unavailability of streamflow data, I decide to set year 2005 as the benchmark. It is based on the fact that in 2005 dry conditions reached a historic level of severity in some parts of Illinois (including McHenry County), and ranked as the one of the three most severe droughts in 112 years of record in Illinois (McHenry County Hazard Mitigation Planning Committee, 2010). As a result, groundwater withdrawal reached a record of 2.636 billion cubic feet in 2005 and these withdrawals of groundwater cast significant impact on stream flow.

3.3.2 Water Price and Price Elasticity of Demand

Water price data are always difficult to obtain, in that there is almost no uniform pricing system in the public water supply sector. Typically, water facilities have a fixed service charge as well as a commodity charge based on the quantity of water consumed. Water rates can be generic, increasing or decreasing across "blocks" of water consumed by households. Surveys are usually sent out to different municipal water supply facilities to collect related water rate information. In this chapter, data on water rates for each individual municipality as of August 2008 were provided by Chicago Metropolitan Agency for Planning. With no available price information for 2005, it is reasonable to use water rates in 2008 as an alternative, since water rates usually undergo very minor changes across years. Water rates for McHenry

County communities are tabulated in Table 3.1.

Table 3.1: Water Rate and Annual Pumpage for Residential Use by Each Municipality in McHenry County

Municipality	Water Price (dollars per thousand cubic feet)	Total Annual Pumpage for Residential Use in 2005 (thousand cubic feet per year)
ALGONQUIN	21.39	18,897
CARY	26.48	100,215
CRYSTAL LAKE	17.13	273,545
FOX RIVER GROVE	17.13	26,110
HARVARD	24.83	53,661
HEBRON	5.91	4,801
HUNTLEY	18.10	45,185
LAKE IN THE HILLS	25.28	128,633
LAKEWOOD	24.68	12,959
MARENGO	23.26	29,510
MC HENRY	26.11	143,634
WONDER LAKE	26.18	16,238
RICHMOND	23.34	8,064
JOHNSBURG	23.49	26,569
WOODSTOCK	10.62	150,365

The municipality-wide estimation of elasticity for residential water use in McHenry County is also not available. Many empirical studies have been devoted to examine demand for water and the own-price elasticity of water. Classic work includes Howe (1982), Nieswiadomy (1982), Hewitt and Hanemann (1995), and Renwick and Green (2000). Several applications of using meta-analysis to identify important factors explaining the variation in estimated price elasticity of residential water demand have been discussed in the literature. Espey *et al*. (1997) found that approximately 90% of reported price elasticities of demand for residential water use were in the range between -0.75 and 0, with a mean price elasticity of -0.51. Another meta-analysis by Dalhuisen *et al*. (2003) reported a mean price elasticity of demand of -0.41 and a median price elasticity of -0.35. Based upon the information above, the residential water price elasticity of demand was set as -0.40. The discussion of sensitivity of elasticities in McHenry County is available in the study by Wan *et al.* (2011). Although it is unavoidable that elasticity is deemed to change under different time and price settings along the water demand curve, adjustments to elasticity are usually difficult to make in a reliable manner. According to Jenkins *et al.* (2003), if deliveries remain within the price range of estimated elasticity, economic loss estimated using constant elasticity is a reasonable approximation.

3.3.3 Features of Wells, Pumping Rate and Aquifer Characteristics in McHenry County

In 2005, there were 100 wells that pumped groundwater to provide daily residential water use for 20 municipalities. However, water rates were unavailable for 13 wells from 5 municipalities (*Island Lake*, *Lakemoor*, *Union*, *Holiday Hills*, and *Prairie Grove*). Thus I decided to use the remaining 87 wells for my analysis. I assigned the municipal water rates to the wells located in that municipality. Data on pumping rate, capacity, geographical location of wells and hydrogeological parameters (specific yield or storativity S_n and transmissivity T_n) are from Meyer *et al.* (2009). Unconfined aquifers frequently have a direct hydraulic connection to rivers, lakes, streams, or other surface water bodies. They have high storativities (typically called specific yield, ranging from 0.1 to 0.3) and release water from storage by the mechanism of actually draining the pores of the aquifer, releasing relatively large amounts of water (Fetter, 1988). In contrast, confined aquifers have low storativity values (much less than 0.01), and they store water using the mechanisms of aquifer matrix expansion and the compressibility of water, which are typically both quite small quantities. Hence, based upon the previous work (Meyer *et al.*, 2009) on the study region around McHenry County, I assume that stream depletion mostly corresponds to the unconfined aquifers and specific yield is used for S_n in equation (3-1). The distance from an individual well to its nearest stream is calculated in ArcGIS through the "proximity" function in "analytical tools". Table 3.2 below tabulates the summary statistics of the variables I collected.

Table 3.2: Statistics of the Variables (Sample Size: n = 87 wells)

Variable	Max	Min	Mean	Standard Deviation
U_n: pumping rate, 1,000 cubic feet per day	124.15	0.01	32.70	28.30
a_n: distance between well and nearest stream, feet	13368.93	148.00	4042.67	3025.15
S_n: specific yield of the aquifer	0.15	0.15	0.15	0.00
T_n: transmissivity of the aquifer, square feet per day	24305.26	250.05	4738.60	4613.20
P_n: price of water , $ per 1,000 cubic feet	26.48	5.91	21.17	5.57

3.3.4 Calibration of Return Flow Coefficient

Township shapefiles in McHenry County were downloaded from the Tiger/Line Shapefile Database at the U.S. Census website. In figure 3.3, I present the geographical visualization of well locations, streams, WWTPs, municipalities and townships in McHenry County. At municipal level, *Crystal Lake* has the highest number (18) of wells used for public water supply, followed by *McHenry* (12 wells), *Cary* (10 wells) and *Woodstock* (8 wells). Table 3.1 shows the total yearly groundwater withdrawal for residential use by each municipality in 2005. *Crystal Lake*, *Woodstock* and *McHenry* rank as the top 3 municipalities for their annual groundwater withdrawal for residential use.

According to the U.S. EPA, there are 16 WWTPs located in McHenry County (see figure 3.3). Data on the locations and daily effluent discharged from WWTPs are retrieved from Clean Watersheds Needs Survey (CWNS) at the EPA website. The survey was conducted every four years, and the data from the 2004 survey were used in this study because it is the closest one to the benchmark year 2005. Based upon the 2004 data, the total daily discharge from all WWTPs was 1.905 million cubic feet and the daily groundwater withdrawal at the entire county-level was 2.46 million cubic feet. Therefore, the return flow coefficient (see equation 3-3) is 77.42%.

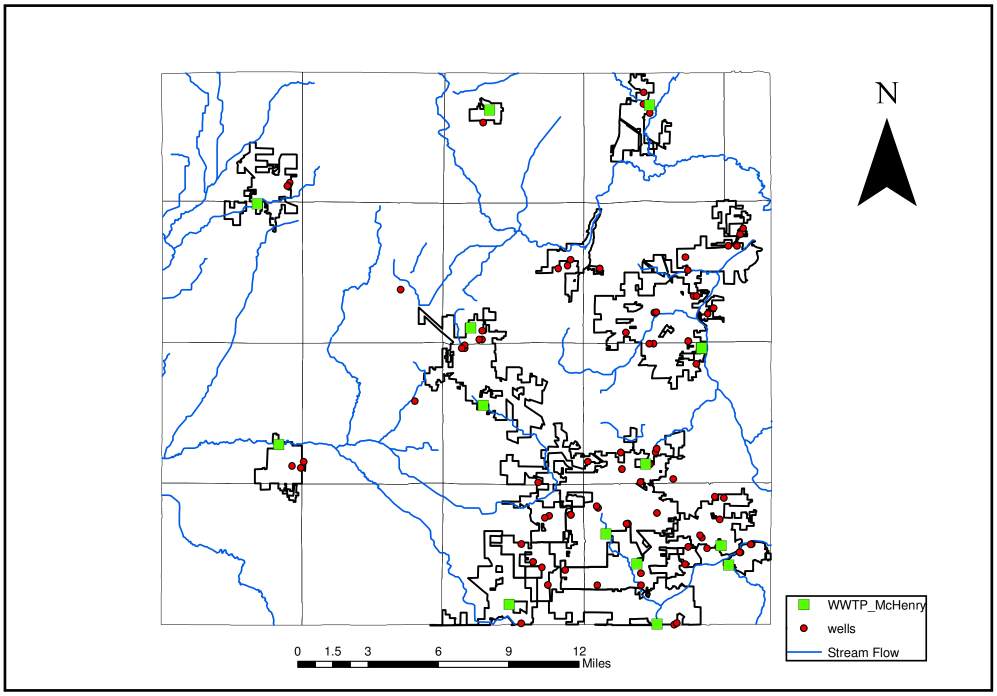

Figure 3.3: Locations of Streams, Pumping Wells and Wastewater Treatment Plants (WWTPs) in McHenry County

3.4 Estimation Results

3.4.1 County Level Analysis

The solution to the multi-objective estimation problem is not a single unique set of parameters but consists of the Pareto set of solutions (non-dominated solutions), according to the various tradeoffs between economic welfare loss and the amount of streamflow to be restored. Using Equations (3-8) – (3-11), the optimization algorithm NSGA-II assigns a pumping rate for each well to produce the Pareto frontier. The optimization is for all 87 wells in the county. In this hypothetical case, my concern is the overall economic welfare at the aggregated county level assuming that there is a centralized planning regime across municipalities.

In figure 3.4, each point represents a Pareto solution of 87 well pumping rates, and these points are equally valid from a multi-objective point of view. It shows the tradeoffs between the amount of streamflow depletion and the associated economic welfare loss due to restriction of pumping rates. When using the NSGA-II search algorithm, the population size and generation size were both set to 1000, and final optimal solutions were nicely distributed along the Pareto frontier. The Pareto frontier allows decision makers or water resource planners to identify solutions based on consequences of the whole set of relative preferences.

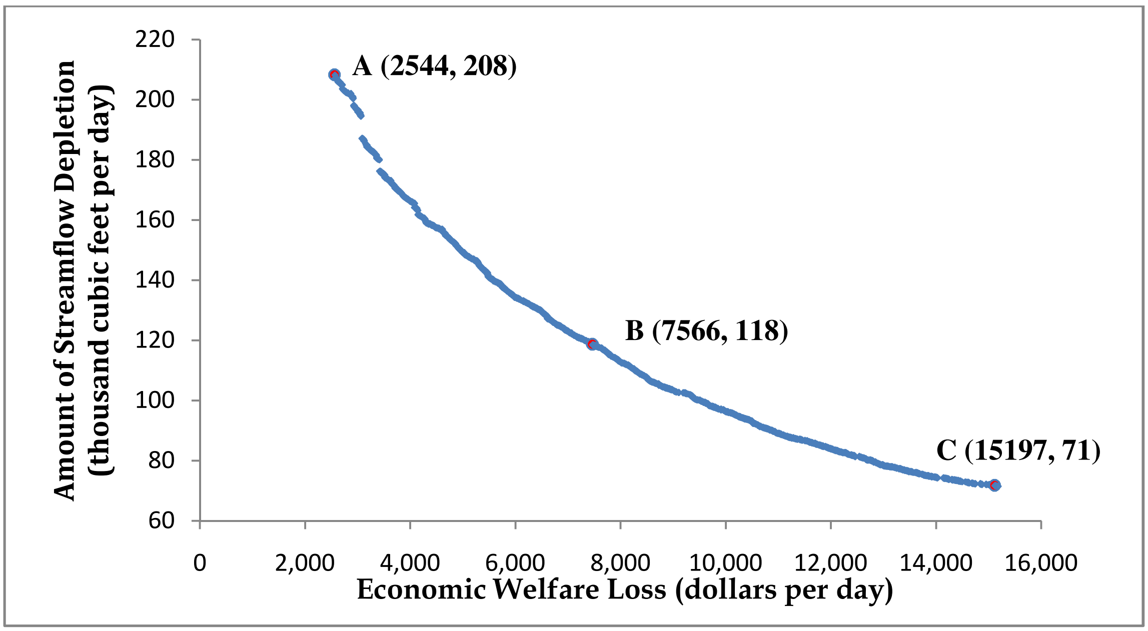

Figure 3.4: The Pareto Frontier Showing the Tradeoffs between Streamflow Depletion and Economic Welfare Loss at McHenry County Level

The shape of the curve indicates that as the amount of streamflow depletion decreases the economic welfare loss increases. More specifically, three search results (two extremes and one middle point) on the Pareto frontier are specified. Point A determines the highest available streamflow depletion level (208 thousand cubic feet per day), while enduring the minimal amount of economic welfare loss (2,544 dollars per day). Point C determines the lowest available streamflow depletion level (71 thousand cubic feet per day), at the expenses of the highest amount of economic welfare loss (15,197 dollars per day). The range of streamflow depletion for decision makers to choose is between 71 and 208 thousand cubic feet (per day), with associated economic welfare costs between 2,544 and 15,197 dollars (per day). Point B shows an intermediate situation with streamflow depletion level at 118 thousand cubic feet (per day) and economic welfare loss of 7,566 dollars (per day).

Identification of three solutions (Point A, B, and C) allows decision makers to further explore the pattern of assigned optimal pumping rates for each well. To compare the spatially heterogeneous effect toward groundwater withdrawal, I summarize the total daily pumpage at each municipality under each scenario in table 3.3. When demand to restore streamflow is low, as in scenario A, the optimal pumping generated by the model at each municipality is just slightly less than its original non-restricted pumping level in 2005. As the demand for preserving streamflow increases, the restriction imposed on the original pumping intensifies. From scenario B to C, daily optimal groundwater withdrawal at each municipality drops in order to lower the amount of stream depletion caused by groundwater pumping, while maintaining the least economic welfare loss. In municipalities, such as *Fox River Grove* and *Harvard*, the solutions of optimal pumping generated by the model under both scenarios almost reach zero. On the other hand, *Cary* and *Wonder Lake* are assigned higher pumping rates from scenario A to scenario B, while the rest of the municipalities are asked to reduce pumping to preserve streamflow. The reason why such a solution occurs is because that *Cary* and *Wonder Lake* have very high water prices (see table 3.1) and restrictions of water supply produce a much higher economic welfare loss. Therefore, to maintain a low economic welfare loss at the entire county level, the decision makers might have to give preferences on water withdrawal allocation to certain municipalities.

Table 3.3: Comparison of Optimal Groundwater Pumping in Three Scenarios and Original Level (Unit: thousand cubic feet per day)

	Point A	Point B	Point C	Original Pumping
ALGONQUIN	49.98	45.28	38.17	51.77
CARY	269.89	273.88	181.32	274.56
CRYSTAL LAKE	738.88	704.73	575.27	749.44
FOX RIVER GROVE	59.91	0.94	0.51	71.53
HARVARD	98.16	0.00	0.11	147.02
HEBRON	8.97	8.28	11.07	13.15
HUNTLEY	122.59	59.38	57.36	123.80
LAKE IN THE HILLS	344.17	319.96	277.83	352.42
LAKEWOOD	33.98	15.06	10.83	35.50
MARENGO	80.67	76.68	33.41	80.85
MC HENRY	363.49	332.76	201.51	393.52
WONDER LAKE	40.82	42.31	32.44	44.49
RICHMOND	18.69	11.87	5.87	22.09
JOHNSBURG	71.02	36.20	22.31	72.79
WOODSTOCK	351.02	172.97	133.64	411.96
TOTAL	2652.23	2100.30	1581.65	2844.89

3.4.2 Municipality Level Analysis

In the previous part, the analytical results are derived assuming that each municipality will honor the entire county's overall economic loss as their primary interests and would collaborate with each other without incurring administrative costs. However, such a concept does not consider the specific incentives of participating municipalities. If I respect the autonomy of each municipality, their decisions on restoring streamflow should be based on the economic welfare losses occurred to the municipalities themselves instead of the entire county. Therefore, I relax the previous assumption and explore the costs of lowering streamflow depletion for each individual community. For Equations (3-8) – (3-11), I no longer aggregate each well *n* toward the entire county level, and instead I examine the wells located at each specific municipality and aggregate the associated economic welfare loss and stream depletion toward that municipality. Correspondingly, the model will produce the Pareto frontier to show the tradeoffs between stream depletion and economic welfare loss for each individual municipal unit. Figures 3.5 shows the Pareto frontiers for all identified 15 municipalities.

Each municipality in McHenry County faces a unique tradeoff situation between streamflow depletion and welfare loss. Municipalities, such as *Huntley*, *Hebron*, *Richmond*, *Lakewood*, and *Algonquin*, show very low capacity for depleting streamflow, mostly due to the fact that in these municipalities only one or two wells are in operation and/or under specific hydrogeological conditions (e.g. low transsimisivity, long distance between the wells and the stream) that the associated streamflow depletion is fairly small. In turn, their Pareto frontiers are all located at the bottom left corner of the graph. *Woodstock*, *McHenry*, *Crystal Lake*, *Cary*, and *Harvard*, on the other hand, are the 5 municipalities with the highest capacities to deplete streamflow and have relatively more impact on the overall streamflow depletion situation at the county level. The slope of the Pareto frontier curve indicates the "shadow price" of the streamflow. For instance, *Woodstock* shows a relatively steep Pareto frontier curve, which suggests that for one unit reduction of streamflow depletion (in other words, one unit of streamflow restoration), the associated economic welfare loss is less. Interestingly, in contrast to the other municipalities, the Pareto frontier of *Lake in the Hills* shows a very flat shape in the tail part. It happens because four out of seven wells in operation at *Lake in the Hills* are located far away from the stream, and according to equation (3-1) their pumping produces very small impact toward streamflow depletion. When restricting the pumping from these four wells, the associated economic welfare loss increases due to lack of water supply, but this does not remarkably restore the streamflow. Hence, these Pareto frontiers together provide decision makers with valuable information on the range of tradeoffs between streamflow depletion and economic welfare loss associated with the unique pumping decisions in each municipality.

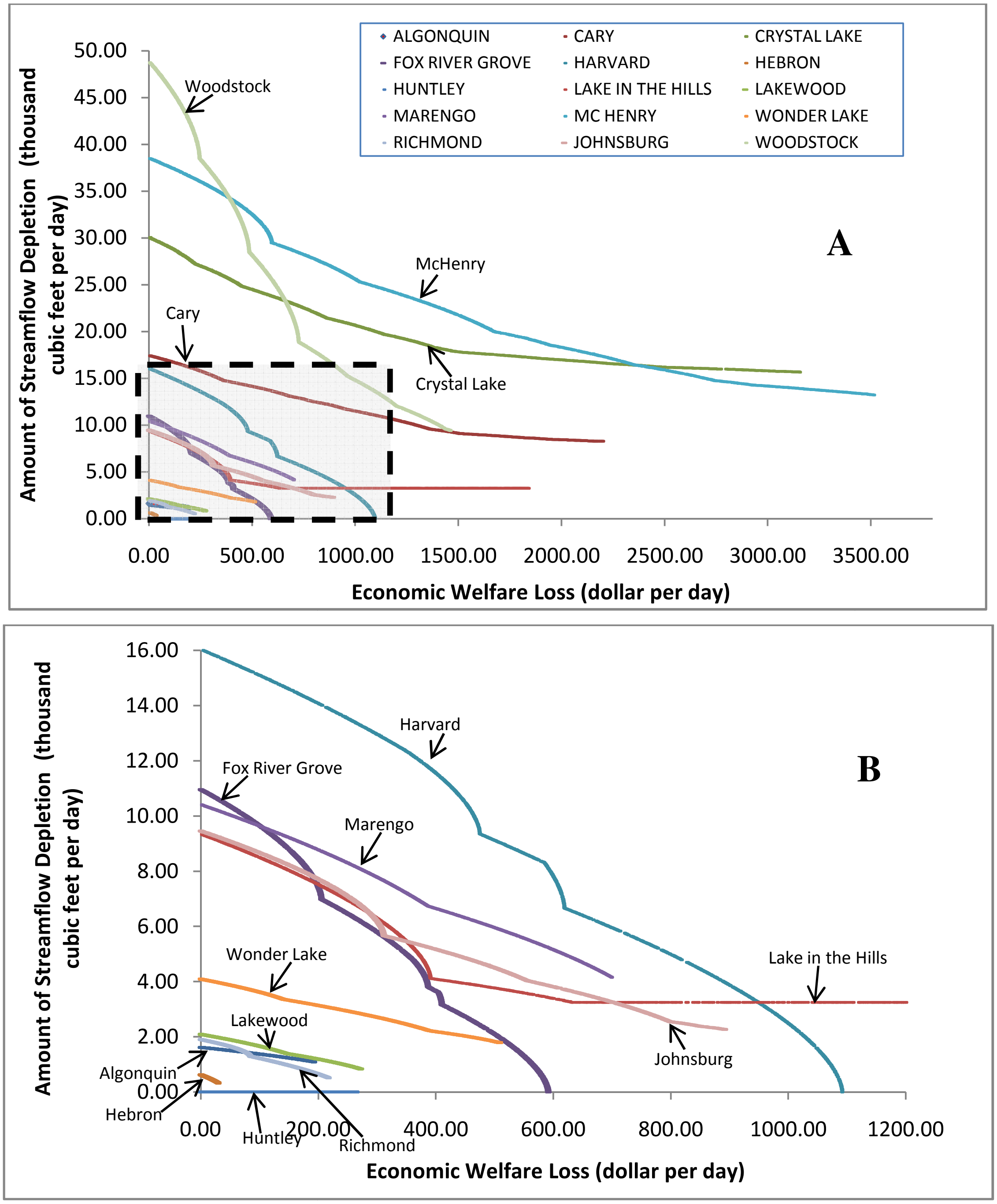

Figure 3.5: The Pareto Frontiers Showing the Tradeoffs between Streamflow Depletion and Economic Welfare Loss for Each Municipality.
A: All Municipalities in McHenry County
B. Municipalities with Low Capacity for Streamflow Restoration

3.5 Discussion

It is noticeable that some municipal Pareto frontiers have a piece-wise shape (such as *Woodstock*, *Harvard*, *Johnsburg* and *Fox River Grove*), while no such feature is observed from the county-wide Pareto frontier. This happens when homogenous price rates are assigned to each well located in one municipality. For one unit less groundwater withdrawal, according to equation (3-7), the associated economic welfare loss at the municipal level will be the same no matter which well's pumping will be under restriction. Acknowledging this further proves that when constraining one unit of groundwater withdrawal (associated with equal welfare loss), reducing the pumping from the well with the highest depletion capacity will restore a maximal amount of streamflow. Under this rationale, the optimal solution along the Pareto frontier (from left to right) will start with restraining pumping from the well with the highest depletion capacity until it reaches its lower bound, and then it moves to the well with the second highest depletion capacity, and goes on till the last well with the least depletion capacity reaches its lower bound. If two consecutive wells along the rank have a remarkable magnitude difference of their depletion capacity, the tradeoffs between restoring streamflow and economic welfare loss will undergo a significant change, and that contributes to the piecewise shape of the Pareto frontier curve in aforementioned municipalities. This finding confirms the importance of spatial planning in producing optimal solutions for groundwater resource management. The well-specific depletion factor $erfc\left(\sqrt{\frac{a_n^2 S_n}{4T_n t}}\right)$, or in other words the hydrogeological factors (distance to streams, specific yield and transmissivity) becomes an important information criterion and suggest a rule of thumb for decision makers to prioritize the preference towards wells in terms of using pumping restrictions to restore streamflow. Given that water rates are the same across all the wells in a municipality, the rule of thumb is to reduce pumpage of the well with the highest stream depletion factor until the streamflow goal is met or the well's lower bound is reached, and then continue this process with the well having the next highest depletion factor.

At an aggregated county level, the optimal solutions can no longer start with the wells having the highest depletion capacity, mainly because the water rates vary across municipalities and economic welfare loss produced by one less unit of groundwater withdrawal is not constant across wells. The rule of thumb no longer holds at the county scale, and the Pareto frontier for McHenry County gets smooth and no obvious step-wise shape is observed.

Another concern is the sensitivity on how the setup of lower bound (*LB*) - 0, 40%, 50%, and 60% - affects the shape of the Pareto frontier. Does it also contribute to the piecewise shape we observe? In order to answer both questions, I remove the lower bound condition in equation (3-10) and just require the restricted pumping rate R_n to be no less than zero. Figure 3.6 shows the comparison of two Pareto frontiers for *Woodstock*: the blue curve is the Pareto frontier with lower bound imposed, while the red curve is the Pareto frontier without lower bound. In the case of no LBs, no restriction is imposed on the minimal pumping of each well, and hence the Pareto frontier without LBs can reach the x-axis. It means all wells are shut down, no streamflow depletion occurs, and the welfare loss achieves the maximum.

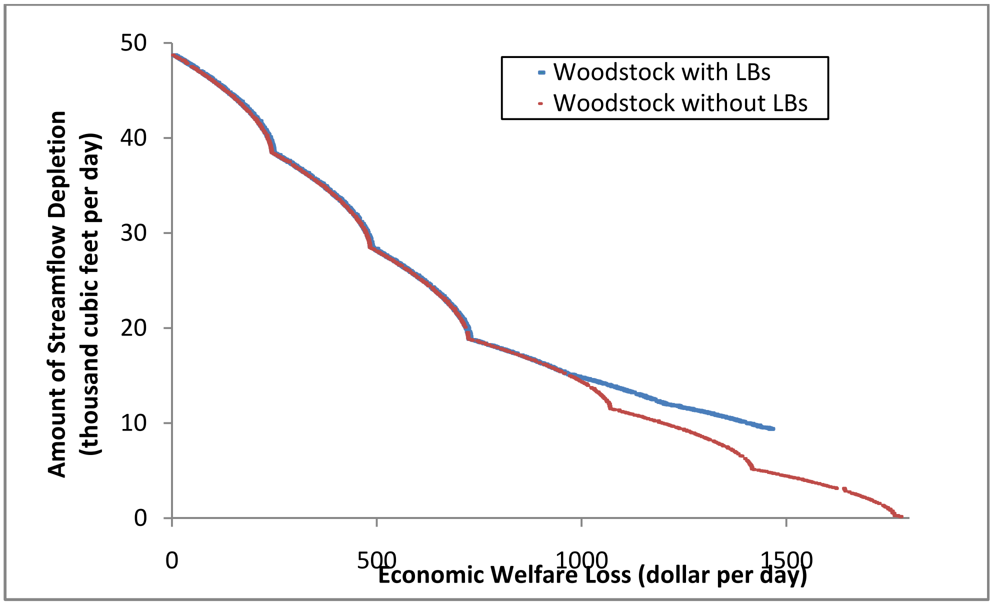

Figure 3.6: Comparison of Pareto Frontiers in Woodstock Municipality for Sensitivity Test of Lower Bounds

The rule of thumb still holds at *Woodstock* for the optimal solutions that comprise the Pareto frontier. Under *LB*s condition, four out of eight operating wells in Woodstock have stream depletion factors in the highest 25% percentile and their pumping could be reduced to zero, while others would not. In the no *LB*s case, all eight wells can have their pumping rate reduced to zero regardless of their depletion capacity. After comparing the two frontiers, I see that the two Pareto frontiers coincide for the first half range (from left to right), simply because these curves from both cases represent optimal solutions that start with the wells of highest depletion capacity and the solutions moves until their pumping reaches zero. After passing the mid-range of the curve, the two Pareto frontiers start to diverge, and the one with *LB*s get flattened off while the one without *LB*s continues with piecewise curvature until it reaches the x-axis. Divergence of the two frontiers is mostly created by the fact that with the *LB*s condition, a minimum pumping requirement is imposed for some wells (which means they are not required to shut down in any case). Following the rule of thumb, when one well reaches the lower bound the optimal solution will switch to reducing the pumping of next well that has less depletion capacity than the previous one along the rank. For wells with above zero lower bound constraints (40%, 50%, or 60%), their available range for pumping reduction is narrower than the ones without a lower bound constraint (from original pumping rate to zero). Therefore, in case of with *LB*s condition, when optimal solutions reach wells with low rank depletion capacity (which are not required to shut down) and move from restricting one well to another, the possible piece-wise curvature is shortened and becomes less obvious. In that sense, the Pareto frontier for the case of *LB*s constraint gets flattened off at the second half and no longer has a piecewise shape. To sum up, the condition of *LB*s will limit the range of pumping restriction of wells in the solution space, and it will somehow change the tradeoff dynamics shown by the Pareto frontier.

The return flow coefficient I used is 77.42%. In order to test the sensitivity to the return flow coefficient, three alternative values (80%, 70%, and 60%) were chosen for the same optimization modeling at the county level. Figure 3.7 shows the comparison of four Pareto frontiers when different return coefficients are used. The population and generation sizes were also set to 1000 in the GA solutions during the sensitivity tests of three alternative return flow coefficients. The Pareto frontier with 80% return flow coefficient is located fairly close to the one I use (77.72%) in this chapter. However, the Pareto frontiers with 70% and 60% return flow coefficients move outward. That is because, as the return flow coefficient decreases (from 77.72% to 70% to 60%), more irrevocable water loss is observed and a smaller portion of groundwater withdrawal is returned to the stream. In that sense, to increase streamflow stock becomes more expensive, since pumping constraints need to be more stringent.

For the four Pareto Frontiers shown in figure 3.7, there are discontinuities of the optimal solutions at similar x-axis values (approximately between economic welfare loss of $3,063 and $3,088 per day). Such a gap indicates that one extra unit of economic welfare loss can contribute to significant improvement of streamflow depletion level. Although I am not expecting continuous solutions generated by GA under the settings of each return flow coefficient, it still remains interesting for me further explore how such gaps are formed. After looking further into the details of the optimal pumping rates produced in the GA solutions that form the two ends of the gap, it is found that at the beginning of the frontier (top left portion), only moderate reductions of the wells' pumping are imposed which leads to small decreases of the streamflow depletion. However, after the observed break near x-axis = $3,063 ~ $3,088, the optimal solution forming the start of a new portion of the Frontier curve starts to have one well significantly reducing its pumping rates (from 31.47 to 4.33 thousand cubic feet per day). This well is found to be located in an area that is very proximate to a stream (thus with high stream depletion capacity) and has relatively low water rates. The gap was again found near x-axis = $3,412 ~ $3,419 with another well's pumping being remarkably reduced (from 63.92 to almost 0 thousand cubic feet per day). This well also has a high depletion capacity (due to moderate distance to a stream and high transimissivity value). In that sense, significantly reducing the pumping of such a well will bring more benefits in terms of lowering stream depletion level while maintaining relatively low economic welfare loss. Such a finding informs me that there is a "threshold" for this trade-off to occur. For the decision of moderately lowering the pumping rate, the solutions will be located at the very beginning of the Pareto frontier (top left portion before the break). For the decision of making an effort to improve the environmental amenity, it will be necessary to significantly restrict the pumping of certain wells (like the two I found at these breaking points) to return more streamflow to the environment. In that way, decision-makers can evaluate the different portions of the Pareto frontier curve based on the trade-off benefit.

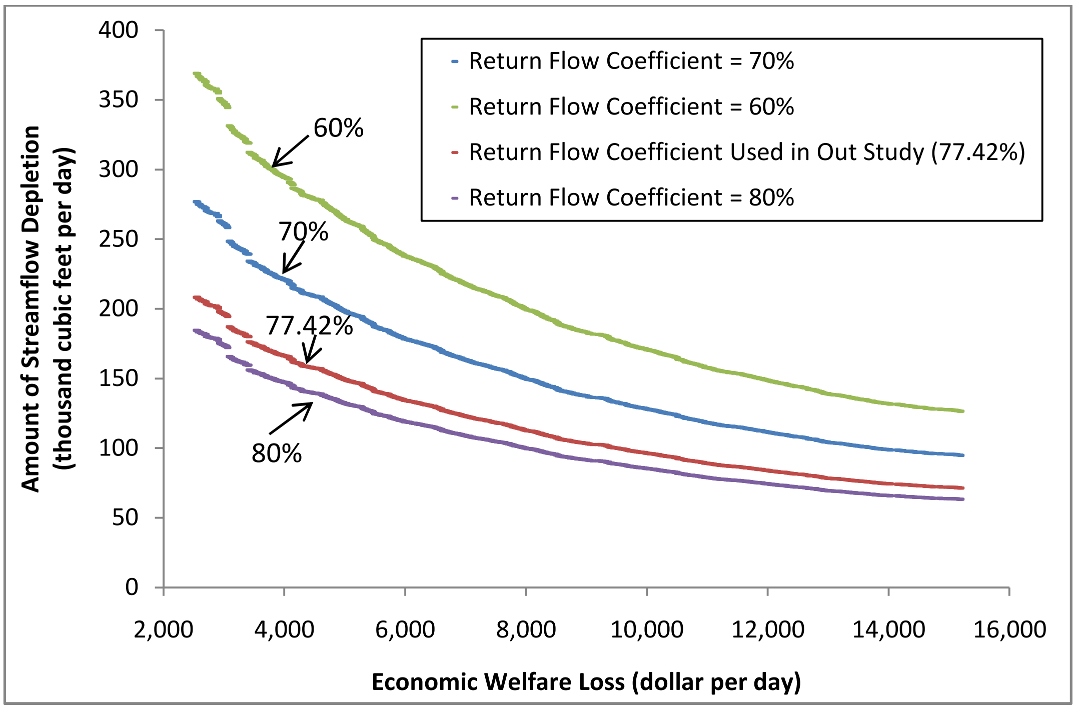

Figure 3.7: Pareto Frontiers under Different Return Flow Coefficients

3.6 Conclusion

This chapter employs the evolutionary algorithms for multi-objective optimization to gain insights on the tradeoffs between stream restoration and economic welfare loss in a hypothetical case study based upon the data available from the McHenry County, Illinois. The decision instruments are based on the groundwater pumping rate of each well, and the study demonstrates that there is a great potential to alleviate the stream depletion problem through optimizing pumping strategies. The NSGA-II algorithm adopted in this chapter provides the decision maker or water resource planner with Pareto frontiers showing a wide range of practical alternatives to manage groundwater resources. By taking advantage of the posteriori approach, the decision makers or water resource planners can choose a preferred solution along the Pareto frontier at a later stage instead of specifying preferences to any of the objectives at the estimation stage.

The shadow price of streamflow found in this study shows a strong spatial heterogeneity, since economic and environmental relations are highly dependent on the spatial locations between wells and streams, hydrogeological characteristics, and water rates. A wide range of regulatory and non-regulatory tools can be tailored to the needs, resources, and political climate of different regional scales. Innovative regulatory tools such as flexible zoning, conservation cluster zoning, and urban growth boundaries can steer groundwater use toward appropriate areas and away from environmentally sensitive ones. At the municipal scale, the rule of thumb serves as a useful guidance for planners and policy makers to effectively select wells to restrain pumping to preserve maximum amount of streamflow. A trade process on streamflow depletion units might be further designed among municipalities to ultimately achieve an economic equilibrium.

The major limitations of the proposed economic-hydrologic-integrated optimization framework arise from several factors. First, the price elasticity of demand is borrowed from the literature and assumed to be constant in the entire county area. In reality, residential users' preferences toward water might vary according to their income level, living habit, housing situations, and so forth. The modeling outcomes will be more informative if better information on elasticities becomes available. Second, streamflow depletion might be one of the negative impacts resulting from excessive pumping and there are other dimensions of environmental amenities deterioration for further exploration, such as water quality, and stream ecology. Issues discussed in this chapter might not suffice to deliver a complete sustainable water resource planning but a proof of concept and guidance on data collection. These factors of environmental concerns can be captured in future studies. Third, this chapter of study is static and short-term based, in that it only considers a bench mark year of 2005. Optimization models can be further improved to produce planning solutions under the long-term conditions of additional wells, change of water prices, dynamics of pumping rates, etc.

To conclude, the merit of this chapter is to propose a modeling framework to determine the optimal spatial and temporal allocation of groundwater via regulating pumping rates in a multi-objective setting of hydrologic and economic concerns. It takes specific consideration of the spatial tradeoffs between economic loss due to constrained groundwater supply and restored streamflow conditions. It confirms the idea that effective planning strategies could be adopted to produce cost-effective outcomes to maintain environmental amenities and ecological values in areas like McHenry County.

Chapter 4

INSPECTING THE OCCUPATION-INDUSTRY LINKAGES IN THE REGIONAL WORKFORCE DEVELOPMENT

4.1 Introduction

The relationship between the industry performance and its impact on the economy is one of the continuing analytical and policy challenges. Florida *et al*. (2008) noted that if you ask a typical person on the street about what drives economic development, his/her answer is probably the jobs created by some booming industries. It is true that well performing industries usually have the potential to stimulate workforce growth and promote new firm spin-offs. However, one shortcoming of industry-based approaches (Tiebout, 1956; Krugman, 1983), as typically practiced, is that policy makers have not focused as much attention on the supply (input) side of industrial production, especially the role of human capital or special skill-mixes in enhancing the productivity and profitability of firms.

Following the pioneer work by Piore and Sabel (1984) who highlighted the role of innovation and human skills as the drivers of regional productivity growth, the concept of treating regional economic development as a function of its concentration of new ideas, innovation, knowledge spillovers and division of labor has been echoed by many other scholars (Lucas, 1988; Saxenian, 1994; Storper, 1997; Mathur, 1999; Scott, 2000; Florida, 2002; Glaeser and Saiz, 2004). The theoretical breakthrough occurred when Romer (1990, 1994) proposed the idea of "endogenous growth theory" in contrast to Solow's (1956) model that treated technology as exogenous in the economic growth process. Romer's (1986, 1987, 1990) theory is based on the premise that economic growth is an endogenous outcome of the economic system and technological advance comes from things that people do, with an emphasis on the linkage among technology, human capital, and economic growth.

The conventional approach that uses educational attainment to account for human capital (Rauch, 1993; Berry and Glaeser, 2005) does not elaborate on how demographic characteristics of a population (e.g. years of education) contribute to a regional economy, and also does not provide an appropriate metric to measure innovative skills and entrepreneurial potential. Even four decades ago, Mincer (1974) recognized that schooling was an incomplete specification of human capital in terms of wage inequalities. Hence, Thompson and Thompson (1987, 1994) proposed a complementary "occupational-function" approach to study local human resources and this idea has been further developed by Feser (2003), Markusen (2004 & 2008), Koo (2005), and Florida *et al.* (2008) to become a potentially more robust measure of utilized skill – that is, how human talent or capability is absorbed by and used in the economy. Following their perspectives, occupation is the mechanism through which education is converted into skill and labor productivity, and becomes part of the signaling and the codification of labor specialization in the economic production process (Sweeney, 2004).

Therefore, understanding the workforce dynamics and economic development potential requires investigating occupations (what people do) as well as industries (where people work) (King *et al.*, 2010). This chapter intends to explore the linkages between occupations and industries in the context of regional workforce development. The findings will provide crucial insights for manpower planners in adjusting training systems to changing economic conditions and to meet the demand for different kinds of skills. In section 4.2, the linkage between occupations and industries is elaborated and an occupation-industry hybrid approach is introduced as an alternative methodology to gain insights into regional workforce dynamics. Section 4.3 establishes a modeling framework to calibrate the magnitude of growth signals of industrial and occupational forces in workforce development dynamics. Also in this section, empirical tests of the model are provided and the results are interpreted by comparing them with earlier findings. Section 4.4 uses a hypothetical industry extraction method to test the sensitivity of occupational demand and identifies the "key industrial sectors" that have high capacity for generating jobs through multiplier effects. The final section 4.5 summarizes and concludes with some policy implications.

4.2 The Occupation-Industry Linkages and Manpower Planning

Since labor demand is derived from economic production, the structural composition of regional occupational groups is the result of constant market selection processes of particular skill mixes needed to maintain competitive and efficient production. In this sense, the mix of executives and managers, scientists and engineers, and skilled, semi-skilled and unskilled workers becomes the enduring legacy of a local workforce. Occupational analysis informs the supply side of regional productivity and growth, and reveals what type of work is taking place and how the varieties of work are linked with one another (Markusen *et al.*, 2004; Currid and Stolarick, 2010).

In addition to industrial performance, well functioning occupational groups become another force to promote regional workforce growth. Geographical concentration of firms in the same or related industries creates a pooled market for workers with similar skills. This reduces the uncertainties about manpower and raises the availability of knowledge and skills. Occupation-based human capital thus enriches the host locality by attracting more industries interested in taking advantage of the skilled labor force, since proximity to people increases the access to the ideas of those people. As global market integration enhances the inter-connections and competitions among places, spatial heterogeneous distributions of labor and scattered firm functionalities have been commonly witnessed. Places gradually distinguish themselves through the functions they perform – financial centers, manufacturing headquarters, cultural and educational leaders, and so forth (Duranton and Puga, 2005; Markusen and Schrock, 2006). Therefore, the growth effect of certain industries may not equally apply to all staffed occupations, but instead some occupations might outperform others and thus grow faster.

After acknowledging the intricate linkage of industry and occupation in their functional roles of promoting workforce growth, manpower planners face the challenge of exploring labor growth signals from both industries and occupations in an interrelated and simultaneous framework. The opportunity emerges when occupation-by-industry workforce matrices become available across different time periods. An occupation-by-industry workforce matrix helps classify the existing workforce into two dimensions and offers a uniquely rich representation of workforce structure. The rows of the matrix represent the specified occupational groups, and the columns of the table represent the industrial sectors. Each cell of the table represents the number employed in one specific occupation-industrial combination. For one specific year, each column of the occupation-by-industry workforce matrix shows the distribution of total employment in this industrial sector across identified occupational groups, thus providing an "industry staffing pattern". Similarly, each row of the occupation-by-industry workforce matrix explains how one occupational group is allocated among different industrial sectors, yielding an "occupation allocation pattern". Analyzing occupation-by-industry tables across time provides insights into the underlying process of structural changes in the past. Certain occupational groups or industrial sectors are likely to produce strong agglomerative effects that are reinforcing (Thompson and Thompson, 1986). Insights into the dynamics of occupation-by-industry employment compositions offer a fresh new perspective to examine and disassemble the functions performed by industrial sectors and occupational groups to jointly promote regional workforce development.

In the United States, job training and workforce development entered the public policy field in the early 20th century, with the federal government as the prime innovator and financial supporter. In 1973 Congress consolidated several of these programs into the Comprehensive Employment and Training Act (CETA) that was later replaced by the Job Training and Partnership Act (JTPA) in the early 1980s. JTPA relies heavily on private business sector involvement through private industry and "state coordinating councils” with overall planning of programs and funds. In similar fashion to the federal efforts, individual states operate their own forms of skill training programs. Effective planning techniques are crucial in providing information that enables training programs to adjust to changing economic conditions to meet the demand for different kinds of occupational skills (Adams *et al.*, 1992).

4.3 Detecting Signals of the Occupational and Industrial Growth Forces

4.3.1 Motivation and Model Formulation

Under an integrated global economy, manpower planners or policy makers need to sense and respond quickly to the occupation needs produced by changes in the demand for goods and services, by the adoption of new technologies, and by the emergence of new industries. There has been an increasing literature (Stolarick and Florida, 2006; Barbour and Markusen, 2007; Currid and Stolarick, 2010) exploring possible ways to link performance of occupational groups with regional economic development. Frequently studied are arts, design, business, and information technology occupations. However, few prior studies developed a holistic perspective to examine the occupational and industrial composition, and subsequent dynamic changes of the regional workforce structure. In order to address these limitations, this chapter proposes consideration the occupation-by-industry composition of the entire regional labor force and extends usually static analyses to a more dynamic context.

As previously discussed, the growth of the workforce could result from the increased demand from a booming industrial sector or it could also be caused by well functioning occupation-based human capital base that yields job opportunities for their own specific occupational group across different industrial sectors. To identify/measure the forces across time from both dimensions, occupation-by-industry workforce tables at two time periods are compiled. A “bi-causative matrix” modeling approach is then adapted to model the forces generating the changes of two occupation-industry workforce matrices over time. [1]

The idea of using a causative matrix was initially proposed by Lipstein (1968) to transform one transition probability matrix at time *t* to the next period *t+1* in a stochastic process to improve Markovian predictions.

$$P_t C = P_{t+1} \tag{4-1}$$

[1] The word “causative” commonly used in input-output studies no longer carries its literal meaning in the context of workforce development studies.

When P_t is nonsingular, C can be derived by

$$C = P_t^{-1} P_{t+1} \tag{4-2}$$

In equations (4-1) and (4-2), a temporally constant matrix C is used to summarize the nature of inter-period change of the two matrices and C is referred to as a causative matrix. Rogerson and Plane (1984) further developed an alternative specification to equation (4-1), in that matrices are not necessarily communicative with respect to multiplication.

$$CP_t = P_{t+1} \tag{4-3}$$

and $C = P_{t+1} P_t^{-1}$ (4-4)

Rogerson and Plane treat equations (4-1) and (4-3) as different views on system change, and denote the causative matrix in equation (4-1) as the right causative matrix, C^R, and the causative matrix in equation (4-3) as the left causative matrix, C^L. In order to capture a mixture of effects denoted by C^L and C^R, Jackson *et al*. (1990) proposed a doubly causative model:

$$C^L P_t C^R = P_{t+1} \tag{4-5}$$

and $p_{ij,t+1} = \sum_{k=1}^{n} \sum_{l=1}^{n} c_{il}^L \, p_{lk,t} c_{kj}^R$ (4-6)

In equation (4-6), the number of unknowns to be estimated from both left and right causative matrices surge to $2n^2$ and render it impossible to find viable solutions. The interpretations of each cell in the left and right matrices become more complex as well. A useful way to handle this problem is to simplify these two causative matrices into two diagonal matrices and the total number of unknowns to be estimated reduces to $2n$ (de Mesnard, 2000). Such a bi-causative matrices approach has been widely used in input-output studies (de Mesnard, 1990 & 1997; Jackson *et al*., 1990; Domingues *et al*., 2002) to interpret structural change in a regional economy. In this chapter, the structure of causative formulation is adapted to model the driving forces and inter-relationships of occupational groups and industrial sectors during the change of regional labor force systems in two time periods.

In equation (4-7), M_{t_1} and M_{t_2} are two occupation-by-industry regional workforce matrices compiled for time t_1 and t_2. There are p occupational groups and q industrial sectors. C^O and C^I are the two diagonal matrices indicating the forces from the occupation side and industrial side that jointly render the changes from matrix M_{t_1} to M_{t_2}. Under matrices multiplication, employment at occupation i and industry j at time t_1, m_{ij,t_1}, is multiplied by c_i^O and c_j^I to become the actual observation of employment at occupation i and industry j at time t_2, m_{ij,t_2}. (See equation 4-9.) c_i^O and c_j^I denote the forces from occupation i and industry *j*.

$$C^O M_{t_1} C^I = M_{t_2} \tag{4-7}$$

Expressed in matrix form,

$$\begin{bmatrix} c_1^O & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_p^O \end{bmatrix} \begin{bmatrix} m_{11,t_1} & \cdots & m_{1q,t_1} \\ \vdots & \cdots & \vdots \\ m_{p1,t_1} & \cdots & m_{pq,t_1} \end{bmatrix} \begin{bmatrix} c_1^I & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_q^I \end{bmatrix} = \begin{bmatrix} m_{11,t_2} & \cdots & m_{1q,t_2} \\ \vdots & \cdots & \vdots \\ m_{p1,t_2} & \cdots & m_{pq,t_2} \end{bmatrix} \tag{4-8}$$

and $c_i^O m_{ij,t_1} c_j^I = m_{ij,t_2}$ (4-9)

With the availability of occupation-by-industry workforce matrices for two time periods at a certain regional scale, unknowns of the diagonal elements in c_i^O and c_j^I in equation (4-8) could be estimated through an optimization process. Equation (4-10) yields the objective of such an optimization process, which is obtained through minimization of the Sum of Squares (SS) of the differences between $m_{ij,t2}$ and $c_i^O m_{ij,t1} c_j^I$ (de Mesnard, 2000):

$$SS = \sum_{i=1}^{m} \sum_{j=1}^{n} [(m_{ij,t2} - c_i^O m_{ij,t1} c_j^I)]^2 \tag{4-10}$$

The results are following expressions:

$$c_i^O = \frac{\sum_{j=1}^{q} m_{ij,t2} m_{ij,t1} c_j^I}{\sum_{j=1}^{q} (m_{ij,t1})^2 (c_j^I)^2}, \text{ for all } i \tag{4-11}$$

$$c_j^I = \frac{\sum_{j=1}^{p} m_{ij,t2} m_{ij,t1} c_i^O}{\sum_{j=1}^{q} (m_{ij,t1})^2 (c_i^O)^2}, \text{ for all } j \tag{4-12}$$

In the model formulation above M_{t_1} and M_{t_2} do not need to be square and they could be rectangular ($p \neq q$). This is a real advantage, "because transforming a naturally rectangular matrix into a square matrix requires additional hypothesis" (de Mesnard, 2000). Interpretations of the estimated C^O and C^I are derived after comparing them with the identity matrix (I). If $C^O = C^I = \text{I}$, then $c_i^O = 1$, $c_j^I = 1$, and $m_{ij,t_1} = m_{ij,t_2}$; the number of people employed at industry *j* holding occupation *i* remains the same from time t_1 to t_2. Employment at the intersection of occupation *i* and industry *j* are under non-changing/neutral forces from two sides. Elements from C^O and C^I matrices, if different from elements of an identity matrix, then indicate forces underlying structural changes inside the regional workforce.

$$\begin{bmatrix} c_1^O & 0 & 0 \\ 0 & c_2^O & 0 \\ 0 & 0 & c_3^O \end{bmatrix} \begin{bmatrix} m_{11,t_1} & m_{12,t_1} & m_{13,t_1} \\ m_{21,t_1} & m_{22,t_1} & m_{23,t_1} \\ m_{31,t_1} & m_{32,t_1} & m_{33,t_1} \end{bmatrix} \begin{bmatrix} c_1^I & 0 & 0 \\ 0 & c_2^I & 0 \\ 0 & 0 & c_3^I \end{bmatrix} = \begin{bmatrix} c_1^O m_{11,t_2} c_1^I & c_1^O m_{12,t_1} c_2^I & c_1^O m_{13,t_1} c_3^I \\ c_2^O m_{21,t_1} c_1^I & c_2^O m_{22,t_1} c_2^I & c_2^O m_{23,t_1} c_3^I \\ c_3^O m_{31,t_1} c_1^I & c_3^O m_{32,t_1} c_2^I & c_3^O m_{33,t_1} c_3^I \end{bmatrix} \tag{4-13}$$

The C^O matrix can be interpreted as a "matrix of occupational forces". The element c_i^O affects equally all terms of *row i*. (For illustration purpose, a simplified 3 × 3 occupation-by-industry workforce model is constructed in equation 4-13.) If element c_i^O is larger than 1, the employment for occupation i across all industrial sectors is multiplied by a value larger than 1 leading to positive growth. In turn, they will be negatively affected if the element c_i^O is less than 1. The C^I matrix can be interpreted as a "matrix of industrial forces". The element c_j^I affects equally all terms of *column* j in matrix M_{t1}. If element c_j^I is larger than 1, the employment for industry j across all occupations is positively affected, while they will be negatively affected if the element c_j^I is less than 1. In general, $C^O$and C^I matrices provide signaling information on fabrication or transformation effects jointly imposed by occupational groups and industrial sectors on regional workforce growth.

Findings on the signaling from "occupational forces" can be compared with recent empirical analyses (Florida, 2002; Markusen, 2004; McGranahan and Wojan, 2007) on human capital or occupation-based theories for economic development. One well-known occupation-mix approach is from Florida's (2002) "creative class theory". According to Florida, creative occupations are defined as those in which individuals "engage in complex problem solving that involves a great deal of independent judgment and requires high levels of education or human capital". Markusen (2004) proposed targeting occupations as well as industries in formulating economic development strategies, including the identification of "key occupations" (Markusen, 2004) where there are "high levels of 'capturability' (proxied by uneven distributions across regions and localities), high levels of absolute and expected growth, high levels of connectivity and cross-fertilization across industries, and high levels of self-employment and entrepreneurship."

Although Florida devoted a great deal of effort to solidify the concept of creative class (or occupational skill) as an important input for regional development, he still left the identified creative classes as exogenous variables in his empirical analysis (Florida *et al.*, 2008). It has been argued that the distinction between human capital theory and the creative class thesis was left ambiguous both theoretically and empirically (McGranahan and Wojan, 2007). Florida's classification might be challenged, since functions performed by these "creative classes" in promoting a regional economy might vary significantly from region to region, and it is not always a tenable idea to draw a generic conclusion. The criteria for choosing "key occupations" proposed by Markusen (2004) are not always precise and are open for different interpretations, such as "connectivity". McGranahan and Wojan (2007) re-examined Florida's creative class thesis in the context of rural areas and argued that the analysis by Florida does not critically assess the construct validity of the creative class measure. Hence, the merit of the modeling approach in this chapter lies in the fact that it takes a holistic approach to examine the dynamic changes of regional workforce and it considers both occupational and industrial forces simultaneously without predetermining the "creativeness" of certain occupations.

4.3.2 Data & Analysis

Occupation-by-industry workforce matrices are compiled at both the national and state level for 2005 and 2008, using data from the Public Use Micro-data Sample (PUMS) files of the American Community Survey (ACS). The PUMS contains an approximately one-percent sample record of the entire population and provides information related to the surveyed people on their occupations and industries. Only one job (with occupation and industry information) is reported for a single worker and thus excludes the possibility of reporting multiple jobs. PUMS files can help cross-classify the number employed under different occupation-industry combinations.

Occupational groups and industrial sectors are specified according to the Census occupation (Standard Occupational Classification, SOC) and 2-digit NAICS industry codes, as listed in table 4.1. All records of the surveyed workers in each state's PUMS file are taken into account to construct the occupation-by-industry workforce table. In the process of occupation-by-industry workforce table compilation, the weights, assigned to individual samples in survey design and post-survey work, are considered. (Note that, O23, the military specific occupations, is excluded in that such occupations are only surveyed in sector I20 (*Federal, state, and local government*) and too many zeros showing in other industrial sectors in the table complicates the analysis.)

Once the occupation-by-industry workforce matrixes for all 50 states (Washington D.C. and Puerto Rico are excluded) as well as the entire U.S. are prepared, analyses are conducted in the manner proposed in equation (4-7) - (4-12). GAMS (General Algebraic Modeling System) is used to accomplish the non-linear optimization to generate solutions for C^O and C^I - more specifically, CONOPT, a nonlinear programming solver, was employed to solve the optimization problem, due to its recognized strength in solving models with a large degree of nonlinearity. It needs to be noted that both upper and lower bounds of the elements in C^O and C^I (i.e. 2.0 and 0.5, respectively) were imposed to control possible diverging estimates (outliers) in a few industrial sectors and occupational groups.

Table 4.1: Descriptions of Occupation Code and Industry Code

Occupation Code	Description	Industry Code	Description
O1	Management occupations	I1	Agriculture, Forestry, Fishing, and Hunting
O2	Business and financial operations occupations	I2	Mining
O3	Computer and mathematical occupations	I3	Utilities
O4	Architecture and engineering occupations	I4	Construction
O5	Life, physical, and social science occupations	I5	Manufacturing
O6	Community and social services occupations	I6	Wholesale Trade
O7	Legal occupations	I7	Retail Trade
O8	Education, training, and library occupations	I8	Transportation and Warehousing
O9	Arts, design, entertainment, sports, and media occupations	I9	Information
O10	Healthcare practitioners and technical occupations	I10	Finance and Insurance
O11	Healthcare support occupations	I11	Real Estate and Rental and Leasing
O12	Protective service occupations	I12	Professional, Scientific, and Technical Services
O13	Food preparation and serving related occupations	I13	Management of Companies and Enterprises
O14	Building and grounds cleaning and maintenance occupations	I14	Administrative and Support and Waste Management and Remediation Services
O15	Personal care and service occupations	I15	Educational Services
O16	Sales and related occupations	I16	Health Care and Social Assistance
O17	Office and administrative support occupations	I17	Arts, Entertainment, and Recreation
O18	Farming, fishing, and forestry occupations	I18	Accommodation and Food Service
O19	Construction and extraction occupations	I19	Other Services (except Public Administration)
O20	Installation, maintenance, and repair occupations	I20	Federal, State, and Local Government
O21	Production occupations		
O22	Transportation and material moving occupations		

4.3.3 Analyses of Results

Table 4.2 tabulates the result of the model estimation on occupation-by-industry workforce changes between 2005 and 2008 in the U.S. As noted earlier, a large magnitude of elements in C^O or C^I (over 1.0) indicates an underlying strong signaling on growth in the occupational groups or industrial sectors respectively. From the estimation (table 4.2), the top 5 groups from the occupation side with strong growth forces are: O18 (*Farming, fishing, and forestry occupations*), O1 (*Management occupations*), O2 (*Business and financial operations occupations*), O13 (*Food preparation and serving related occupations*), and O3 (*Computer and mathematical occupations*). By the same token, the top 5 groups from industry side (table 4.2) are I2 (*Mining*), I13 (*Management of companies and enterprises*), I14 (*Administrative and support and waste management and remediation services*), I3 (*Utilities*) and I4 (*Construction*). However, I1 (*Agriculture, forestry, fishing, and hunting*) and I6 (*Wholesale trade*) seem to lack forces in promoting workforce growth.

Table 4.2: Estimation of Occupational Forces C^O and Industrial Forces C^I for US (National)

Occupation Code	Estimates of C^o	Industry Code	Estimates of C^I
O1	1.071	I1	0.833
O2	1.069	I2	1.415
O3	1.064	I3	1.093
O4	1.013	I4	1.088
O5	1.033	I5	0.971
O6	1.050	I6	0.881
O7	0.993	I7	1.040
O8	0.996	I8	1.032
O9	1.011	I9	0.968
O10	1.022	I10	0.968
O11	1.027	I11	0.951
O12	1.050	I12	1.053
O13	1.067	I13	1.126
O14	1.013	I14	1.098
O15	1.055	I15	1.060
O16	1.013	I16	1.046
O17	0.999	I17	1.085
O18	1.272	I18	1.031
O19	0.925	I19	1.015
O20	0.982	I20	0.997
O21	0.956		
O22	1.040		

Judging from the estimated magnitudes of the forces, it is interesting to notice some mixed results that O18 (*Farming, fishing, and forestry occupations*) shows an underlying strong growth momentum, whereas I1 (*Agriculture, forestry, fishing, and hunting industry*) follows an opposite pattern. For O18 & I1 combination (farming, fishing, and forestry occupations at agriculture, forestry, fishing, and hunting industry), during 2005 – 2008, the number employed increased from 1.18 million to 1.26 million. The low estimate from the industrial side indicates I1 (*Agriculture, forestry, fishing, and hunting industry*) itself may not generate enough "demand power" to create incentives for the workforce growth across all occupations. However, there are significant increases of employment in O18 (*Farming, fishing, and forestry occupations*) across several industry sectors. For instance, O18 (*Farming, fishing, and forestry occupations*) at the I16 (*Health care and social assistance industry sector*) increased from 464 to 1554. It is possible that the skill-linkage associated with O18 (*Farming, fishing, and forestry occupations*) helps inject vigor into the labor market as the I16 (*Health care and social assistance industry sector*) gains its growing momentum. More farming related occupational workers are hired to support a wide range of services (especially food products, gardening, logging, and etc.) needed by the health care industry. In that sense, one should be cautious about ascribing negative contributions for agricultural occupations based on the agricultural sectors' lack of growth. A recent study by Currid and Stolarick (2010) identifies similar findings about the information system/information technology (IS/IT) occupation and industry in Los Angeles. They found "a substantial number of those IS/IT occupations are working in categorically non-IS/IT industries" and the real strength of Los Angeles' IS/IT skill/knowledge cluster is embedded in other industries other than the IS/IT industry. Another possible explanation is the sampling error or bias associated with American Community Survey. Even though I've carefully considered the number of people surveyed and their representing weights in the labor market, it is still possible that in 2005, farming occupational workers employed by health care industries were less extensively surveyed as opposed to the case in 2008, and the increase of 464 to 1554 could produce very sensitive results to the "bi-causative" estimation system, not only because their small numeric values but also the rate of change ratio.

The model is extended to 50 individual states in the U. S. (excluding Washington D. C. and Puerto Rico) to estimate the state-specific occupational and industrial forces. State-level analyses provide additional insights toward the workforce dynamics at smaller regional scales. Among others, the analyses reveal both the magnitudes of growth forces from occupations and industries in each state and their spatial heterogeneity. Generally, among the estimates, a considerable degree of variance is found, whereas some occupations and industries show consistent positive influence in most states. In the following, attention will be directed to the implications of these outcomes. (See the table 4.3 and 4.4 for the detailed state-level estimations.) Since the data are retrieved from a survey, there are missing values in the occupation-by-industry tables. Those missing values do not necessarily mean no worker is employed at such occupation-by-industry categories, but instead probably the survey sample is not comprehensive enough to cover such information. Nevertheless, without better data available, the entry for this scenario is recorded as zero in the occupation-by-industry workforce tables, which results in N.A. (not available) estimates in the modeling outcomes. In order not to bring possible bias to complex the conclusion towards the interpretations of occupational and industrial forces, these N.A. values are replaced by a neutral value

"1".

One example is provided in figure 4.1 where each occupation or industry is represented as a dot on the X (Mean) – Y (Standard Deviation) coordinates. This figure is intended to place the industries/occupations with strong forces to promote workforce growth (cross-state averages of estimated elements of C^O or C^I are larger than 1) on the right side, and the industries/occupations with high variations across states in the upper portion (cross-state variances of the elements of C^O or C^I are larger).

Table 4.3: Estimation of Occupational Forces C^O for 50 States (2005 – 2008)

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	O12	O13	O14	O15	O16	O17	O18	O19	O20	O21	O22
AK	0.728	0.711	0.893	0.86	0.993	1.6	0.85	1.193	1.373	1.216	1.435	1.191	1.531	1.5	1.546	1.266	1.148	1.3	0.819	1.298	1.038	1.253
AL	1.083	1.128	1.095	0.932	0.767	1.107	0.817	1.019	1.066	1.148	1.052	1.118	1.034	0.971	1.081	1.088	1.016	1.137	0.896	1.035	0.954	1.103
AR	0.983	1.073	0.776	1.08	0.874	0.813	0.86	1.013	0.833	1.161	1.055	1.013	1.053	0.812	0.934	1.049	0.95	0.88	0.636	0.89	0.9	0.896
AZ	1.032	1.061	1.044	1.007	0.876	1.174	0.968	0.917	0.858	1.036	1.13	1.105	0.958	1.051	1.169	0.972	1.028	0.895	0.831	0.932	0.866	0.955
CA	1.171	1.152	1.281	1.081	1.215	1.126	1.145	1.06	1.194	1.098	1.126	1.222	1.287	1.17	1.309	1.159	1.129	1.317	0.968	1.083	1.073	1.207
CO	1.028	0.984	0.899	0.895	0.856	0.925	0.832	0.914	1.007	0.932	0.862	0.83	1.159	0.995	0.97	0.92	0.909	1.356	0.918	0.926	0.984	1.01
CT	1.237	1.272	1.229	1.039	1.07	1.604	0.886	1.116	1.351	1.319	1.359	1.298	1.28	1.158	1.412	1.186	1.14	0.5	1.013	1.104	1.081	1.291
DE	1.131	0.827	0.622	0.971	0.61	1.02	1.016	1.093	0.597	1.293	1.355	1.256	1.155	1.242	1.828	0.919	1.04	0.5	1.038	1.041	0.822	0.692
FL	1.185	1.233	1.177	1.087	1.156	1.134	1.173	1.125	1.133	1.083	1.141	1.201	1.293	1.171	1.036	1.057	1.122	1.15	1.067	1.093	1.06	1.038
GA	0.957	0.945	0.972	0.907	0.8	0.91	1.015	0.892	0.812	0.904	0.873	0.872	0.899	0.844	0.968	0.973	0.871	1.187	0.933	0.839	0.831	0.929
HI	0.876	1.147	1.125	1.567	1.088	1.022	0.955	1.041	0.963	0.88	1.324	0.848	0.862	0.733	0.985	0.723	0.969	0.693	0.916	0.799	0.708	0.916
IA	1.051	1.105	0.808	1.176	0.862	1.136	1.015	0.837	1.067	1.016	1.013	1.026	1.164	0.885	1.037	0.989	0.979	1.307	0.993	0.847	1.01	1.06
ID	1.492	1.635	2	1.438	1.257	1.382	0.924	1.69	1.742	1.632	1.405	1.412	1.969	1.316	1.519	1.692	1.554	1.624	1.547	1.111	1.547	1.255
IL	1.049	1.053	1.034	1.007	1.067	1.044	0.989	0.948	1.059	1	1.049	0.981	0.851	1.06	1.041	0.984	0.995	1.275	1.021	0.943	0.992	1.051
IN	1.085	1.074	1.102	1.035	0.997	0.947	1.154	0.968	0.866	1.083	1.04	1.147	0.897	0.854	0.892	0.921	0.95	1.331	1.002	1.009	0.87	0.953
KS	1.451	1.746	1.378	1.411	1.403	1.887	1.802	1.8	1.532	1.387	1.465	1.429	1.502	1.412	1.192	1.439	1.37	1.716	1.328	1.211	1.201	1.273
KY	1.062	1.044	1.232	1.004	0.873	0.823	1.081	0.89	1.136	0.958	0.997	1.027	0.961	1.184	0.889	0.96	0.987	1.097	1.038	0.924	0.899	1.048
LA	1.087	0.988	0.752	1.098	0.584	1.227	0.746	0.943	1.239	0.996	1.033	0.98	0.955	0.949	1.003	0.952	0.913	0.679	0.977	0.901	0.905	0.999
MA	0.808	0.756	0.715	0.696	0.776	0.62	0.78	0.691	0.697	0.734	0.781	0.778	0.781	0.688	0.763	0.753	0.735	0.794	0.63	0.683	0.581	0.653
MD	1.138	1.058	1.01	1.132	0.933	0.898	1.089	1.144	1.03	1.01	1.039	1.064	1.103	1.011	0.976	1.043	0.989	1.262	1.013	1.119	0.96	1.059
ME	1.407	1.665	1.489	1.303	1.204	1.146	1.539	1.32	1.293	1.573	1.472	1.74	1.583	1.711	1.341	1.431	1.245	1.7	1.213	1.217	1.506	1.399
MI	1.094	1.075	1.034	1.031	1.019	1.164	0.789	0.997	0.901	1.099	1.144	1.215	1.12	1.018	1.043	1.022	1.016	1.518	0.799	0.994	1.035	1.127
MN	1.287	1.256	1.155	1.204	0.943	1.27	0.995	1.104	1.229	1.036	1.062	1.076	1.192	1.039	1.196	1.187	1.2	1.431	1.106	1.126	1.126	1.263
MO	1.035	1.111	1.098	0.964	1.041	1.194	1.032	1.149	0.849	1.018	0.922	0.975	0.999	0.938	1.148	1.057	0.95	1.388	0.892	1.006	0.896	1.047
MS	1.127	1.069	1.324	1.31	1.118	1.043	0.997	0.976	0.891	1.115	1.127	0.936	0.87	1.063	0.874	0.88	1.027	0.828	0.973	0.909	0.928	0.942
MT	1.116	1.172	0.947	0.955	0.729	0.801	0.883	0.703	0.825	0.779	1.156	1.14	0.649	0.697	1.07	1.067	1.002	1.565	0.786	1.055	0.899	0.911
NC	1.691	1.627	1.687	1.372	1.686	1.644	1.717	1.668	1.659	1.595	1.522	1.358	1.661	1.605	1.654	1.711	1.629	2	1.352	1.707	1.44	1.672
ND	0.671	0.63	0.575	0.85	0.678	0.572	0.618	0.62	0.683	1.192	0.962	2	1.166	1.038	1.109	0.952	0.727	0.981	0.818	0.612	0.5	0.794
NE	1.441	1.432	1.36	1.53	1.205	1.387	1.135	1.249	0.881	1.026	1.061	1.858	1.094	1.11	1.073	1.15	1.191	1.489	1.569	1.212	1.344	1.247
NH	1.421	1.34	1.186	1.1	0.792	1.541	1.014	1.502	1.164	1.129	1.151	1.061	1.099	1.265	1.193	1.197	1.135	1.609	1.124	1.112	1.005	1.131
NJ	1.267	1.224	1.248	1.202	1.261	1.188	1.203	1.299	1.132	1.322	1.091	1.108	1.374	1.305	1.346	1.238	1.161	1.488	1.096	1.134	1.173	1.386
NM	1.006	1.058	0.954	0.83	0.998	0.88	0.864	0.919	0.946	0.872	0.849	1.046	0.833	0.901	0.941	0.965	0.9	1.186	1.006	0.793	1.076	1.048
NV	1.052	1.02	1.129	0.881	0.5	1.005	0.951	1.387	0.942	1.125	1.039	1.09	1.114	1.217	0.972	0.931	0.902	0.901	1.111	0.714	0.94	1.025
NY	1.037	1.125	1.017	1.063	1.004	1.071	0.951	0.964	1.002	1.006	1.047	1.047	1.151	1.071	1.064	0.975	1.01	1.459	0.864	0.977	0.916	1.049
OH	1.04	1.105	1.18	1.048	0.953	1.079	1.067	0.991	0.975	0.974	0.945	1.123	1.145	0.983	1.035	1.048	1.019	1.422	0.986	1.017	1.014	1.075
OK	1.196	1.019	1.142	1.071	1.103	1.076	1.256	1.097	0.937	1.007	0.991	1.002	1.022	0.973	0.869	0.994	1.074	0.892	0.995	0.891	0.966	1.081
OR	1.198	1.322	1.127	1.151	1.316	1.212	1.075	1.127	0.999	1.264	1.077	1.374	1.418	1.248	1.039	1.142	1.161	1.277	1.017	0.947	0.945	1.12
PA	1.057	1.012	1.024	1.032	1.024	1.024	0.886	0.935	1.005	0.99	1.058	1.067	1.066	1.035	1.124	1.02	0.971	1.231	0.91	0.938	0.897	1.086
RI	0.527	0.608	0.529	0.599	0.7	0.68	0.67	0.546	0.594	0.552	0.544	0.598	0.57	0.5	0.634	0.671	0.635	2	0.533	0.628	0.644	0.553
SC	1.152	1.095	1.22	1.211	1.134	1.178	1.167	1.167	0.962	1.172	1.167	1.246	1.171	1.05	1.104	0.979	1.069	0.855	0.908	1.036	1.018	0.961
SD	0.643	0.619	0.502	0.835	0.577	1.145	1.135	1.079	0.677	0.88	1.135	0.961	0.738	0.711	1.172	0.525	0.836	0.691	0.902	0.71	0.693	0.662
TN	1.161	1.15	1.302	1.168	1.046	1.123	1.263	1.01	1.097	1.207	1.095	1.153	1.224	1.14	1.099	1.157	1.052	1.344	0.869	1.009	0.99	1.03
TX	1.056	1.07	0.935	1.052	1.071	1.09	0.905	1.018	0.994	1.014	1.003	1.023	1.021	0.96	1.029	1.011	0.984	1.603	0.976	1.042	1.035	1.016
UT	1.159	1.071	1.168	1.329	0.958	1.134	1.167	1.171	0.961	1.126	1.063	1.106	1.152	1.081	1.065	0.977	0.937	1.341	0.963	0.869	0.973	0.98
VA	1.055	1.081	1.029	0.965	0.959	0.877	0.913	1.051	0.91	0.994	1.144	1.097	1.001	1.078	1.135	1.025	0.963	1.563	0.849	0.985	0.979	1.052
VT	0.898	0.754	0.905	0.633	0.806	1.493	0.881	0.908	0.5	0.84	0.979	1.215	1.015	0.906	0.543	0.826	0.751	1.561	1.11	0.592	0.8	0.852
WA	1.096	1.057	1.012	1.099	1.096	1.075	1.072	1.035	1.134	1.041	0.958	1.036	1.167	1.006	1.048	1.016	1.003	1	0.939	1.019	0.96	0.981
WI	0.606	0.597	0.671	0.543	0.5	0.638	0.611	0.581	0.582	0.632	0.692	0.726	0.628	0.594	0.66	0.615	0.62	0.646	0.563	0.622	0.644	0.69
WV	1.013	0.938	0.867	0.523	0.852	0.969	1.061	0.768	0.696	1.126	0.906	0.718	1.12	0.803	1.051	0.918	1.021	0.547	0.996	0.837	0.949	1.005
WY	1.034	0.532	0.533	0.87	1.497	1.302	1.114	0.967	0.668	0.932	1.333	0.712	0.947	1.028	1.027	1.219	1.119	1.311	0.943	0.892	0.778	0.879

Table 4.4: Estimation of Industrial Forces C^I for 50 States (2005 – 2008)

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20
AK	0.597	1.799	0.722	1.13	1.452	1.263	0.81	0.66	0.5	1.553	1.107	0.926	1	1.16	0.85	0.768	0.791	0.607	0.631	1.169
AL	0.686	0.924	0.92	1.171	0.986	0.76	0.917	1.091	1.032	0.836	0.975	1.025	1.074	1.141	1.074	0.919	0.925	1.055	1.011	1.091
AR	0.873	1.996	1.159	1.467	1.058	0.886	1.055	1.113	1.197	1.006	0.97	1.008	1.036	1.203	0.979	1.002	1.011	1.047	1.136	0.967
AZ	1.036	1.683	1.116	1.223	0.991	0.807	1.189	1.121	0.98	0.974	0.963	1.023	0.895	1.023	1.143	1.07	0.955	1.189	1.204	1.047
CA	0.898	1.362	1.026	1.003	0.846	0.764	0.92	0.933	0.897	0.808	0.808	0.907	0.967	0.909	0.966	0.958	0.886	0.908	0.87	0.88
CO	0.796	1.9	1.349	1.204	1.106	0.957	1.161	1.148	0.985	1.064	1	1.225	0.693	1.264	1.129	1.216	1.231	0.95	1.194	1.287
CT	1.62	0.507	0.842	0.91	0.856	0.674	0.883	0.741	0.613	0.825	0.868	0.883	1.051	0.973	0.912	0.859	0.773	0.891	0.848	0.83
DE	0.931	1	0.827	1.017	1.064	0.992	1.313	0.802	1.434	0.887	0.823	1.224	0.741	0.93	0.937	0.829	0.797	1.003	0.565	1.05
FL	0.913	0.755	1.131	0.892	0.907	0.787	1.017	0.977	0.835	0.886	0.814	0.904	0.835	0.929	0.966	0.983	0.954	0.872	0.969	0.893
GA	0.838	1.206	1.182	1.102	1.16	1.02	1.068	1.153	1.129	1.053	1.034	1.177	1.544	1.372	1.292	1.192	1.48	1.329	1.132	1.224
HI	1.995	1	1.063	1.403	1.019	1.406	1.211	1.071	0.909	1.087	1.032	0.998	1	1.146	0.982	1.058	1.248	1.244	0.939	1.11
IA	0.823	0.908	0.633	0.988	0.995	0.829	1.055	1.069	0.931	1.097	0.993	1.038	0.768	1.257	1.163	1.048	1.061	1.007	1.097	1.073
ID	0.607	2	0.562	0.677	0.687	0.679	0.75	0.946	0.781	0.614	0.51	0.756	0.5	0.695	0.646	0.659	0.68	0.548	0.751	0.652
IL	0.858	1.361	1.166	0.99	0.965	0.885	1.076	1.001	0.882	1.054	0.886	1.062	1.266	1.059	1.064	1.03	1.145	1.166	1.054	0.886
IN	0.801	1.498	0.984	0.986	1.015	0.934	1.087	1.027	1.059	1.014	1.114	1.145	0.974	1.388	1.204	1.067	1.228	1.215	1.137	0.979
KS	0.508	0.979	0.86	0.835	0.768	0.631	0.75	0.741	0.646	0.734	0.743	0.64	2	0.821	0.666	0.734	0.643	0.742	0.694	0.76
KY	0.782	1.009	0.983	0.935	1.014	0.872	1.094	1.017	1.221	0.997	1.017	1.054	1.172	1.064	1.105	1.044	0.877	1.086	1.041	0.894
LA	1.178	1.354	1.422	1.087	1.046	0.928	1.063	0.878	0.961	0.958	1.076	1.088	1.316	1.079	0.983	1.035	1.037	1.03	1.007	0.97
MA	1.449	0.5	1.433	1.587	1.412	1.097	1.415	1.561	1.346	1.385	1.192	1.494	0.869	1.534	1.524	1.474	1.56	1.537	1.498	1.271
MD	0.737	0.587	1.086	1.015	0.882	0.843	0.994	1.029	0.987	0.867	0.989	1.023	0.5	1.02	0.975	1.075	1.178	1.06	0.985	0.95
ME	0.663	2	0.831	0.699	0.619	0.591	0.753	0.652	0.74	0.805	0.571	0.662	0.701	0.869	0.694	0.723	0.587	0.659	0.734	0.652
MI	0.901	1.047	0.959	1.049	0.865	0.804	1.004	1.014	0.923	0.971	0.93	1.042	0.86	1.06	0.985	0.942	1.039	0.95	0.993	0.955
MN	0.721	0.912	0.919	0.861	0.858	0.842	0.923	0.894	0.778	0.813	0.818	0.862	1.035	1.021	0.925	0.969	0.883	0.821	0.91	0.904
MO	0.738	1.435	1.087	1.109	1.028	0.883	0.999	0.854	0.942	1.053	0.9	1.079	1.06	1.276	0.996	1.051	1.287	1.012	0.953	1.011
MS	0.914	1.619	1.084	1.17	0.896	0.853	1.064	1.087	0.762	0.911	1.037	1.063	1	0.952	1.1	0.965	0.934	1.228	0.992	0.933
MT	0.724	1.99	0.906	1.225	1.093	0.83	0.963	1.009	0.613	0.998	1.124	1.156	1	1.448	1.365	1.148	1.162	1.228	1.037	0.882
NC	0.652	0.758	0.732	0.767	0.616	0.615	0.658	0.612	0.599	0.635	0.665	0.684	0.562	0.664	0.692	0.701	0.845	0.641	0.653	0.731
ND	1.387	2	1.051	1.162	1.48	1.301	1.112	1.597	1.024	1.959	0.646	1.737	1	1.895	1.369	1.066	0.505	0.967	0.774	1.561
NE	0.577	0.906	0.952	0.685	0.799	0.705	0.864	0.808	0.848	0.712	1.296	0.902	2	0.901	0.85	0.938	0.83	0.836	0.858	0.68
NH	0.593	2	0.662	0.805	0.887	0.81	0.863	0.822	1.016	0.852	0.769	0.815	1.767	1.075	0.735	0.967	0.986	0.859	0.66	0.817
NJ	0.589	0.727	0.925	0.979	0.759	0.79	0.869	0.758	0.734	0.796	0.762	0.894	1.114	0.894	0.847	0.845	0.837	0.807	0.812	0.808
NM	0.785	1.378	1.155	1.041	1.147	0.742	1.039	1.037	1.097	1.017	0.934	1.065	1	1.341	1.113	1.104	1.192	1.019	1.158	0.96
NV	1.078	1.105	1.121	0.983	1.008	0.939	1.282	1.147	0.918	0.973	0.963	1.066	0.5	1.208	0.979	1.036	0.812	1.186	1.114	1.154
NY	0.811	0.829	1.093	1.139	0.953	0.938	1.084	1.05	0.969	0.944	0.993	1.053	1.419	1.077	1.099	1.04	1.143	0.98	1.055	0.99
OH	0.819	0.88	0.988	0.974	0.925	0.807	0.994	1.013	0.991	0.943	0.843	1.072	0.596	1.1	1.063	1.055	1.054	0.904	1.004	0.867
OK	1.028	1.292	1.028	1.116	0.99	0.773	1.028	0.977	0.872	0.878	0.92	0.93	1.44	0.912	0.892	1.078	1.334	1.044	1	0.99
OR	0.731	1.35	0.811	0.978	0.902	0.681	0.899	1.095	0.85	0.806	0.841	0.97	0.953	0.8	0.927	0.979	1.031	0.824	1.086	0.809
PA	0.841	0.986	1.037	1.13	1.046	0.936	1.056	0.996	1.051	0.975	0.871	1.079	1.195	1.082	1.085	1.017	1.105	0.987	0.983	0.997
RI	0.733	2	1.455	1.892	1.521	1.786	1.566	1.981	1.065	1.547	1.506	1.976	1	1.714	1.963	1.84	1.915	1.881	1.148	1.672
SC	0.948	2	0.67	1.128	0.868	0.783	0.987	1.219	0.949	0.931	1.179	1.069	0.5	1.074	0.951	0.976	0.975	1.059	0.977	0.935
SD	1.512	0.5	0.797	1.352	1.411	1.853	1.775	1.681	1.229	1.602	0.988	1.558	1	1.196	1.003	1.333	1.589	1.396	0.674	1.275
TN	0.895	1.205	1.039	1.176	0.868	0.777	0.918	1.059	0.933	0.961	0.942	0.898	1.072	1.106	1.067	0.937	0.974	0.887	0.892	1.064
TX	0.693	1.499	1.057	1.115	0.979	0.936	1.035	1.121	0.965	1.025	1.057	1.13	0.987	1.159	1.037	1.081	1.187	1.115	0.985	1.041
UT	0.965	1.29	1.067	1.117	0.92	0.951	1.111	1.035	1.023	0.985	1.009	0.989	0.589	1.168	1.042	1.043	1.013	1.08	0.91	0.995
VA	0.79	1.31	1.261	1.164	0.936	0.795	1.009	1.04	0.957	0.965	0.824	1.12	1.908	1.164	1.057	1.061	1.235	1.098	0.98	1.037
VT	1.104	0.5	1.697	0.841	1.038	1.049	1.309	1.644	1.138	0.985	1.681	0.982	1	1.931	1.11	1.29	1.448	1.132	1.162	0.863
WA	0.968	1.51	1.069	1.158	0.97	0.912	1.022	1.057	0.913	0.895	0.971	1.113	0.787	1.088	1.005	1.058	0.881	0.981	0.988	1.058
WI	1.41	1.073	1.922	1.542	1.692	1.286	1.587	1.758	1.577	1.725	1.423	1.845	0.649	1.82	1.72	1.59	1.878	1.593	1.638	1.55
WV	0.938	1.422	0.704	1.074	1.007	1.012	1.139	0.765	0.813	0.89	0.793	1.074	1	1.016	1.168	1.04	1.113	0.995	0.957	1.198
WY	0.596	1.137	0.928	1.209	1.699	1.14	0.89	0.984	1.102	0.81	1.387	0.906	1	1.005	0.925	1.008	0.795	0.924	1.039	1.239

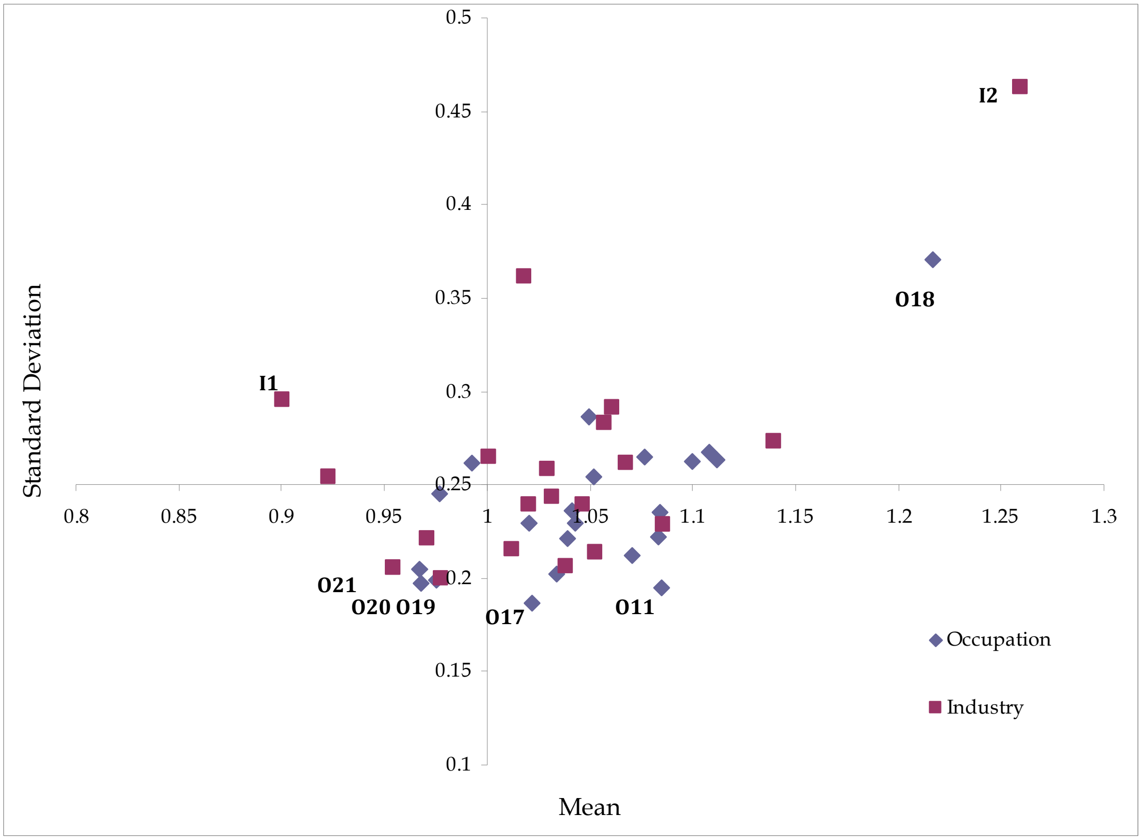

Figure 4.1: Mean and Standard Deviation of the Occupational Forces C^O and Industrial Forces $C^I$for 50 States

First of all, from the occupation side, most of the cross-state averages of the estimates for C^O elements are between 1 and 1.1, indicating a positive force on workforce growth. The ***highest*** cross-state averages of the estimates is for O18 (*Farming, fishing, and forestry occupations*), which is similar to the conclusion the occupational group with the highest estimate at national level. Nevertheless, the ***high*** variance associated with O18 also indicates that such strong force to promote workforce growth does not equally apply to all states. One explanation is that there are spatial heterogeneities about the performance of O18 in regional workforce growth dynamics. Another explanation is that these results are possibly biased and are sensitive due to the quality of American Community Survey data. In terms of variance, the ***lowest*** cross-state variance of the estimates for occupational groups is found for O17 (*Office and administrative support occupations*), indicating that functions performed by this occupational group are mostly stable across states and are less spatially heterogeneous. The combinations of ***low*** estimate of cross-state average with the ***low*** cross-state variance are found at O21 (*Production occupations*), O20 (*Installation, maintenance, and repair occupations*) and O19 (*Construction and extraction occupations*). This finding shows that these occupations exert relatively ***weak*** strength to promote state-level workforce growth, and it is weak almost across all the 50 states with small spatial variances.

Secondly, when it comes to the industrial side, from figure 4.1, most of the cross-state averages of the estimates on the C^I elements are located on the right side of the Y-coordinate, indicating that most industrial sectors cast positive effects on regional workforce growth. The cross-state variances of these estimates, however, are generally larger than cross-state variances from the occupation side. The ***highest*** cross-state average of the estimates is found for I2 (*Mining*), which is also similar to previous findings at the national level. Growing forces from I2 (*Mining*) sector could vary remarkably on a spatial scale, indicated by this sector having the highest cross-state variance. Mining usually requires a relatively high intensity of labor input, and subsequent drastic numerical increase in mining industries in some states is interpreted as sending strong growing forces to the state-level workforce from the model outcomes. The lowest cross-state average of estimates for industrial sectors is found for I1 (*Agriculture, forestry, fishing, and hunting industry*), and it is with extremely high cross-state variance, indicating strong spatial heterogeneity. From table 4.4, the ***lowest*** estimate (0.508) is found in Kansas, and the ***highest*** estimate (1.995) is found in Hawaii.

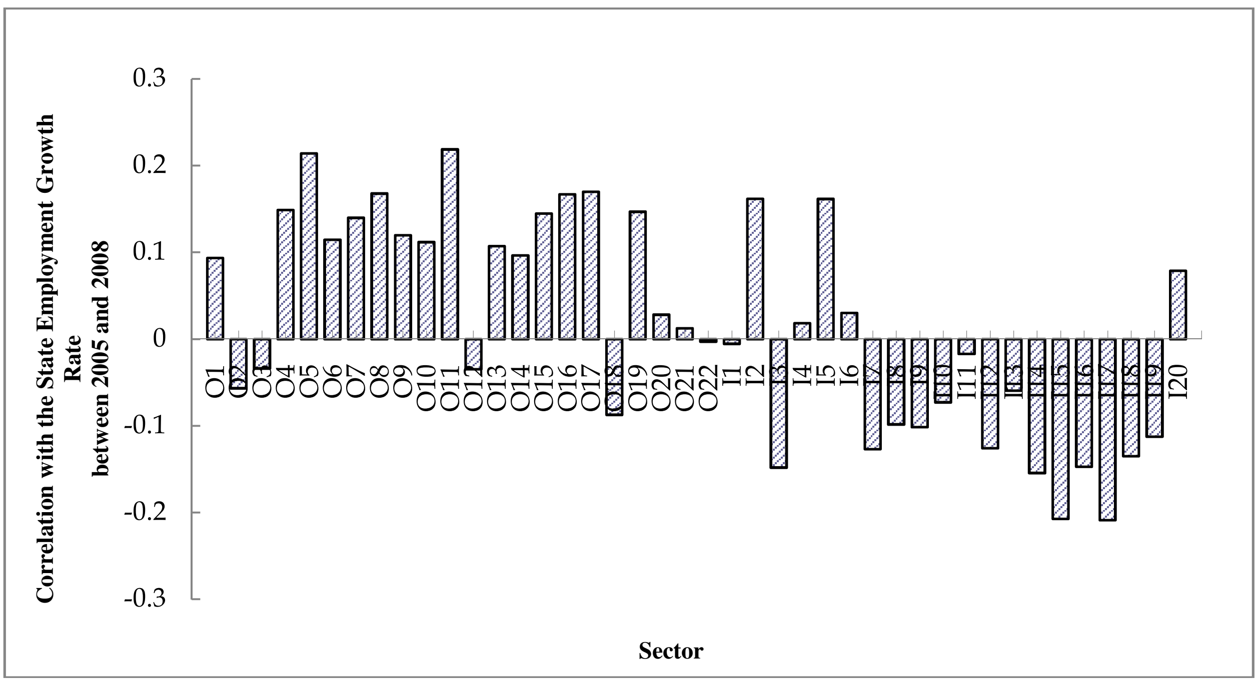

Figure 4.2: Correlation between the estimates of the forces and state employment growth rate

Figure 4.2 shows the correlations between estimated elements of each state-level industrial force or occupational force (C^I or C^O) and the state's total employment growth rate between 2005 and 2008.[2] The intent here is to check if a particular occupation or industry's growth force is highly associated with overall state's employment increase, since Florida (2002) and Markusen (2004) had emphasized the important functions performed by "creative class" and "key occupations", such as engineering, information technology, arts, and entertainment occupations in generating economic growth.

Sectors that show relatively higher positive correlations are mostly from the occupation side. This indicates that certain occupational groups could serve as the key source to inject vigor into the regional labor market. Occupational groups showing high positive correlation with regional labor growth rates are: O11 (*Healthcare support occupations*), O5 (*Life, physical, and social science occupations*) and O17 (*Office and administrative support occupations*). Although the magnitude is relatively small, O9 (*Arts, design, entertainment, sports, and media occupations*) also shows a positive correlation with their state's employment growth.

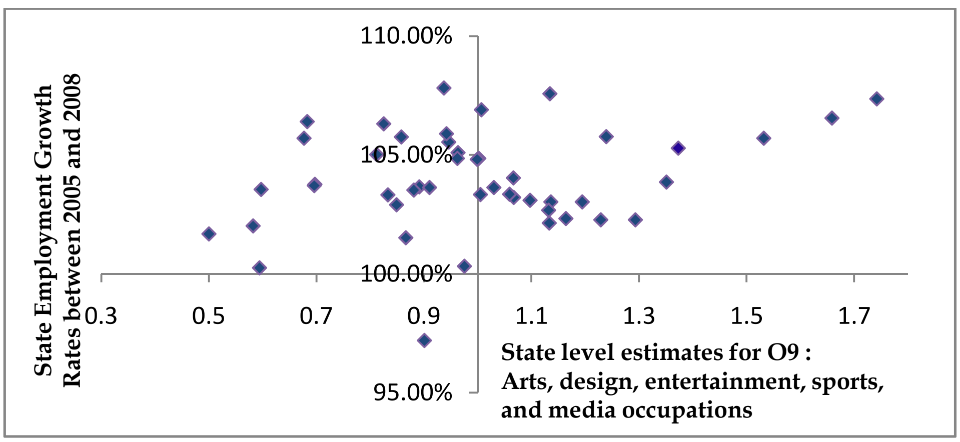

Figure 4.3: Plots of state level estimates for O9 and their employment growth rates (2005-2008)

In figures 4.3, state level estimates of occupational force for O9 (*Arts, design, entertainment, sports, and media occupations*) during 2005 to 2008 are plotted against the state level employment growth rates in the same period. The positive relationships between occupational forces and growth rates are noticeable. However, such strong implications of occupational estimates towards workforce growth are not found among industrial estimates from the model outcomes. For instance, state level estimates of industrial force for I17 (*Arts, entertainment, and recreation*) are plotted against the state level employment growth rates in figure 4.4 and no strong trends are detectable; the correlation between these two variables even turns out to be negative. To conclude based on these observations of the estimates under such an occupation-

[2] To compute the state-level employment growth rate, we used a data series from Bureau of Economic Analysis, because this information may be more accurate than PUMS, ACS in measuring the number of employment.

industry-hybrid approach, estimates of growth forces from the occupation side may serve as a better proxy to understand regional workforce growth.

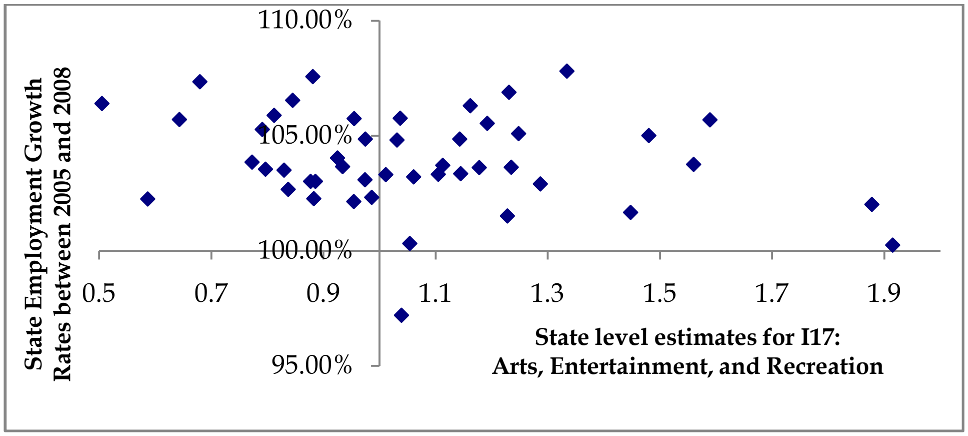

Figure 4.4: Plots of state level estimates for I17 and their state employment growth rates (2005-2008)

4.4 Sensitivity Analysis of Occupational Demand and Identification of Key Industries

Stimulated by the concerns over the loss of key industries, hypothetical extractions are frequently used to measure the inter-industry linkages in an input-output analytical structure (Meller and Marfan, 1981; Dietzenbacher *et al.,* 1993). Assuming that one industry sector *E* is removed from the economy and all intermediate and final requirements of sector *E*'s output are replaced by imports, sector *E* no longer buys or sells any inputs to the remaining sectors, the backward and forward linkages disappear as well. On-going economic activities then have to adjust to this change, and the hypothetical extraction of sector *E* will result in losses in gross output at all sectors on the assumption that there is no spatial substitution of inputs from other economies or technological substitution from other sectors.

Under the industry-occupation framework, the quantity of goods produced by each industry requires labor input through a combination of various occupations. The removal of sector *E* not only eliminates the needs of its own demand for occupations (direct effect), but also reduces the demand of occupations (indirect effect) of the remaining industrial sectors that supply intermediate inputs, directly and indirectly, to sector E. According to Meller and Marfan (1981), the capacity of an industrial sector to generate employment corresponds to the sum of direct and indirect labor requirements. When assuming a constant relationship between output and occupational demand in each industrial sector for a given year, the sensitivity that an industrial sector casts on the overall demand of occupations can be illustrated through this extraction method.

Given the usual input-output system (see equation 4-14), which in this case is a 20-sector economy specified in 2-digit NAICS industry codes (I1 ~ I20),

$$X = AX + Y \tag{4-14}$$

where X is a vector of gross output [20×1]; A is a direct input coefficients matrix [20×20]; Y is a vector of final demand [20×1], the vector of gross output can be expressed as

$$X = (I - A)^{-1}Y = ZY \tag{4-15}$$

where I is a [20×20] identity matrix and Z is the Leontief inverse [20×20].

When a sector E is extracted, all of sector E's linkages to the rest of the economy are eliminated. The input coefficient matrix becomes A^* [19×19], which is obtained by removing from A the row and column of sector E. This captures the comprehensive impact of sector E toward the economy by removing all connections – forward, backward, and internal. The adjusted gross output X^* for the remaining sectors can be measured as:

$$X^* = (I^* - A^*)^{-1}Y^* \tag{4-16}$$

where I^* is a [19×19] identity matrix and Y^* is a vector of final demand [19×1] after removing sector E.

To show the impact on occupational demand due to reduced gross output, an H matrix [22×20] is calibrated to denote the occupational requirement for each industrial sector to produce one unit of gross output, which is obtained by dividing employment in occupation i for industry j (element $m_{i,j}$ in occupation-by-industry workforce matrix M) by the gross output of that industry sector (x_j) in a given year.

$$H = \begin{bmatrix} m_{1,1}/x_1 & \cdots & m_{1,20}/x_{20} \\ \vdots & \vdots & \vdots \\ m_{22,1}/x_1 & \cdots & m_{22,20}/x_{20} \end{bmatrix} \tag{4-17}$$

Therefore, the impact on overall occupational demand across all remaining industries (ΔOD) after extracting sector E can be estimated by comparing the original overall demand for each occupation prior to extraction with the resulting demand for each occupation after extraction.

$$\Delta OD = OD - OD^* = H \cdot X - H^* \cdot X^* = H \cdot (I - A)^{-1}Y - H^* \cdot (I^* - A^*)^{-1}Y^* \tag{4-18}$$

where H^* matrix [22×19] is a matrix of occupational requirement per unit output for the remaining sectors after removing sector E at a given year.

4.4.1 Extraction Analysis at National Level

U.S. national 20-sector input output tables are compiled from BEA Input-Output Accounts for 2005 and 2008. Impacts on occupational demand, by industrial sector, for each of the 20 possible extractions are calculated through equation (4-18). Moreover, the impact on occupational demand (*ΔOD*) due to the hypothetical extraction of industry *E* can be further disaggregated into self-induced effects and nonself-induced effects. The self-induced impact (direct effects) denotes the loss of occupations for the industry *E* itself and the non-self-induced impact (indirect effects) refers to the occupations loss in other industries due to inter-industry linkages.

The fact that the self-induced impact generates high percentage loss of one entire occupational group due to extracting one or several industrial sectors indicates these sectors are the key industries demanding such an occupation. In addition, through indirect effects, the extraction of one industry can indirectly result in a significant portion of employment loss in one or a few occupational groups. Industry of this kind can be considered as the "key industry" to indirectly demand such occupations. The findings on key industries from both direct and indirect effects are summarized in table 4.5. The percentage number in the brackets shows the percentage of loss of the occupational group due to hypothetical extraction of such an industry under either effect. If the direct/indirect effects cast on one occupation group are relatively small and even across each industry that is extracted, it is concluded that no available (N.A.) key industries exist for that specific occupation.

For instance, in terms of direct effects, I1 (*Agriculture, forestry, fishing, and hunting*) is the key industry to demand O18 (*Farming, fishing, and forestry occupations*); I20 (*Federal, state, and local government*) is the key industry to demand O12 (*Protective service occupations*); I9 (*Information*) is the key industry to demand O9 (*Arts, design, entertainment, sports, and media occupations*), and so forth. In other terms, the extraction of industry I1 (*Agriculture, forestry, fishing, and hunting*) will directly lead to 56.8% loss of the O18 (*Farming, fishing, and forestry occupations*), regardless of the non-self-induced impact. The self-induced impacts on O1 (*Management occupations*), O17 (*Office and administrative support occupations*) and O20 (*Installation, maintenance, and repair occupations*) from extractions are relatively small across all industries, suggesting that these occupations are less sensitive to the performance of certain industrial sectors and no serious influential industrial sector(s) exists.

In terms of the non-self-induced impact (indirect effects), key industries can be identified in the same manner. In the same table, it is found that most of the strong impacts through indirect effects are generated by industries I5 (*Manufacturing*) and I20 (*Federal, state, and local government*). For instance, the extraction of I5 will indirectly result in 45.65% loss of the total O18 (*Farming, fishing, and forestry occupations*). In other words, the performance of I5 (*Manufacturing*) has a strong impact on the overall demand of O18 (*Farming, fishing, and forestry occupations*).

Table 4.5: Identification of Key Industries for Occupational Demand through Direct and Indirect Effects in 2005

	Through Direct Effect	Through Indirect Effect
	KEY INDUSTRY	KEY INDUSTRY
O1	N.A.	I5 (*9.0%*), I20 (*7.7%*)
O2	I10 (*22.7%*)	I5 (*10.3%*)
O3	I12 (*31.4%*)	I5 (*12.6%*), I20 (*12.1%*)
O4	I5 (*33.4%*), I12 (*34.9%*)	I5 (*10.2%*), I20 (*12.9%*)
O5	I12 (*26.4%*), I20 (*24.4%*)	I5 (*9.4%*), I20 (*9.4%*)
O6	I16 (*49.2%*)	N.A.
O7	I12 (*58.9%*)	I5 (*12.5%*), I20 (*12.7%*)
O8	I15 (*89.2%*)	I20 (*6.6%*)
O9	I9 (*28.7%*), I12 (*17.5%*)	I20 (*10.6%*)
O10	I16 (*77.5%*)	N.A.
O11	I16 (*86.5%*)	N.A.
O12	I20 (*60.6%*)	N.A.
O13	I18 (*79.2%*)	N.A.
O14	I14 (*36.9%*)	I20 (*10.4%*)
O15	I16 (*31.7%*)	N.A.
O16	I7 (*59.8%*)	N.A.
O17	N.A.	N.A.
O18	I1 (*56.8%*)	I5 (*45.7%*)
O19	I4 (*75.6%*)	N.A.
O20	N.A.	I5 (*8.2%*), I20 (*7.1%*)
O21	I5 (*72.9%*)	I20 (*11.2%*)
O22	I8 (*27.0%*)	I5 (*13.3%*)

To compare the overall extraction effects (direct plus indirect) toward the demand of occupations among individual industrial sectors, figure 4.5 presents the overall percentage loss of each occupation that results from the extraction of six typical industries. The outcomes reveal that when extracting I2 (*Mining*), I3 (*Utilities*) and I17 (*Arts, entertainment, and recreation*), the overall effects have relatively small impacts across all occupational groups. In contrast, extracting I5 (*Manufacturing*), I12 (*Professional, scientific, and technical services*) and I20 (*Federal, state, and local government*) will create significant percentage loss across a broad range of occupational groups, and these industries have stronger linkages that affect overall occupational demand at the national level.

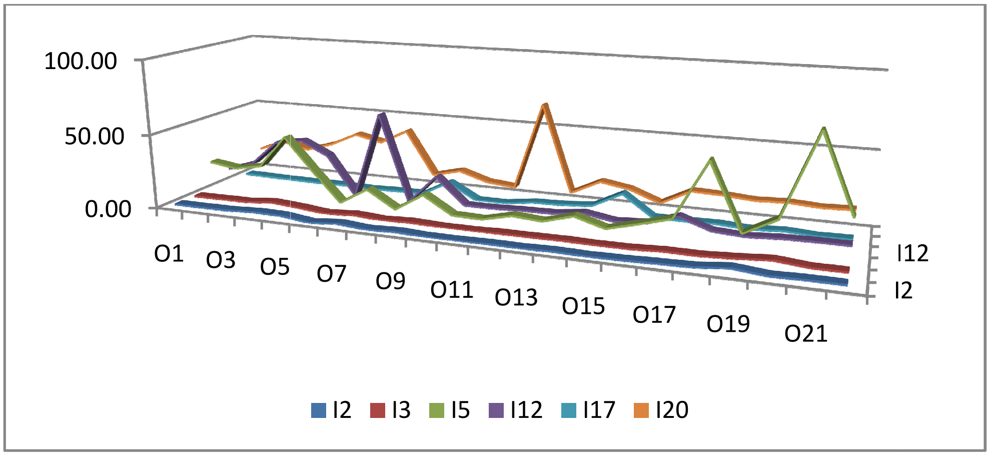

Figure 4.5: Total Effects on Percentage Loss of Occupational Demand due to Hypothetical Extraction of Six Industrial Sectors in 2005

Similar industrial extraction analysis was conducted for 2008. The results on both self-induced and non-self-induced impacts show very similar patterns to those in 2005, revealing that in the short-term (3 years) the inter-linkages between industries and their influences on subsequent demand for occupations remain stable. To save space, the results for 2008 are not included in this chapter and are available upon request.

In order to show the impact on overall employment generation (across all occupations) from industrial sectors at the U.S. national level, two types of criteria can be used in selecting an industry as a "well connected sector". One criterion is a "per output linkage coefficient" that can be obtained through scaling total employment effect (direct plus indirect effect) under the hypothetical extraction process by the gross output (in million dollars) of that industrial sector. The procedure denotes how many jobs in the economy will be lost/gained if one million dollars less/more output is produced by one industrial sector. Another criterion is a "per employment linkage coefficient", and this is calculated by scaling total employment effect (direct plus indirect effect) under hypothetical extraction by the total number of jobs for that industrial sector. This criterion establishes how many jobs in the economy will be lost/created if one job is eliminated from/added to the industrial sector. These linkage coefficients serve a similar purpose as "multipliers" used in input-output analysis.

Table 4.6 shows under both types of criteria for 2005 and 2008 these estimates are mostly stable across 3 years. The high "per employment linkage coefficients" are seen for I11 (*Real estate and rental and leasing*), I1 (*Agriculture, forestry, fishing, and hunting*), I2 (*Mining*), and I3 (*Utilities*), while the high "per output linkage coefficients" are seen for I15 (*Educational services*), I18 (*Accommodation and food services*), I14 (*Waste management and remediation services*), I7 (*Retail trade*), and I16 (*Health care and social assistance*).

Table 4.6: Employment Linkage Coefficient and Output Linkage Coefficient in 2005 and 2008

	Per Employment Linkage Coefficient		Per Output Linkage Coefficient	
	2005	2008	2005	2008
I1	3.31	3.76	4.85	4.31
I2	3.22	3.10	4.15	3.43
I3	3.20	2.94	4.51	3.57
I4	1.79	1.67	10.38	9.88
I5	1.64	1.70	5.03	4.34
I6	1.45	1.47	7.82	7.19
I7	1.20	1.20	15.11	14.85
I8	1.42	1.47	10.25	9.28
I9	1.94	2.08	5.55	5.15
I10	1.59	1.75	4.82	4.70
I11	4.01	3.24	3.51	2.62
I12	1.42	1.39	7.62	6.81
I13	1.53	1.52	7.92	7.27
I14	1.16	1.17	16.68	14.83
I15	1.04	1.04	68.18	58.04
I16	1.28	1.28	14.19	12.87
I17	1.28	1.31	12.93	11.77
I18	1.20	1.20	19.84	18.61
I19	1.43	1.47	10.42	9.86
I20	1.98	2.03	7.24	6.40

Table 4.7: Breakdown of Per Employment Linkage Coefficients across 22-Occupation for Industries Showing Significant Linkage Effects

	I11	I1	I2	I3
O1	0.25	0.14	0.19	0.17
O2	0.27	0.12	0.17	0.17
O3	0.12	0.07	0.10	0.10
O4	0.07	0.08	0.12	0.15
O5	0.03	0.03	0.06	0.05
O6	0.01	0.01	0.01	0.01
O7	0.04	0.02	0.03	0.03
O8	0.03	0.07	0.05	0.03
O9	0.05	0.03	0.03	0.03
O10	0.04	0.02	0.02	0.02
O11	0.02	0.01	0.01	0.01
O12	0.09	0.03	0.03	0.03
O13	0.18	0.06	0.08	0.13
O14	0.26	0.07	0.07	0.07
O15	0.05	0.04	0.03	0.03
O16	0.54	0.19	0.19	0.15
O17	0.98	0.50	0.55	0.64
O18	0.01	0.69	0.01	0.01
O19	0.14	0.09	0.45	0.28
O20	0.25	0.14	0.18	0.38
O21	0.26	0.49	0.43	0.37
O22	0.31	0.44	0.39	0.35
Sum	4.01	3.31	3.22	3.20

Table 4.7 shows the further breakdown of the "per employment linkage coefficient" across the entire range of occupational groups for four industries with high "per employment linkage coefficients" in 2005. For Industrial Sector I11 (*Real estate and rental and leasing*), every 100 jobs lost/created in this sector will generate a total of 401 jobs losses/gains in the economy. Among these 401 jobs, 98 positions will be under O17 (*Office and administrative support occupations*); 54 positions will be under O16 (*Sales and related occupations*); also substantial numbers of positions will be created under O22 (*Transportation and material moving occupations*), O2 (*Business and financial operations occupations*), O21 (*Production occupations*), O14 (*Building and grounds cleaning and maintenance occupations*), O20 (*Installation, maintenance, and repair occupations), and O1 (Management occupations*).

4.4.2 Extraction Analysis at the State Level: Illinois

Due to the difficulty in obtaining state level input-output tables, in this study for illustration purposes, state level analysis on sensitivities of occupational demand due to hypothetical industrial extraction is conducted only for Illinois. The input-output table used for Illinois is compiled from IMPLAN Illinois Industry Transaction Matrix. Although IMPLAN collects state-level input-output information annually, only IMPLAN 2006 is available for me to use. Therefore, in the following analysis, Illinois input-output table in 2006 is used to proxy for the input-output table for Illinois in 2005.

Table 4.8 tabulates the key industries for the demand of each occupational group in Illinois State in 2005 for both direct (self-induced) and indirect (non-self-induced) effects. In terms of direct (self-induced) effects, the identified industries that cast significant impact on the demand of each occupation are similar to those at the national level. It is worth noticing that instead of having no key industries for certain occupations at the national level, in Illinois, I5 (*Manufacturing*) turns out to be the key industry for the O1 (*Management occupations*), I10 (*Finance and insurance*) and I16 (*Health care and social assistance*) for the O17 (*Office and administrative support occupations*), and I19 (*Other services except public administration*) and I5 (*Manufacturing*) for the O20 (*Installation, maintenance, and repair occupations*). When it comes to the indirect effects, Illinois State shows a slightly different pattern of linkages among industries and occupations. I5 (*Manufacturing*) continues to strong influences on the demand of most occupations through input-output linkages, while I20 (*Federal, state, and local government*) no longer has significant indirect impact like it shows at the national level.

Table 4.9 shows the "per employment linkage coefficients" and "per output linkage coefficients" for the Illinois in 2005. In contrast to the national level, the highest per employment linkage coefficients are seen at I13 (*Management of companies and enterprises*) and I2 (*Mining*). The highest per output linkage coefficient is seen at I15 (*Educational services*), which is the same as at the national level. Other industrial sectors with high per output linkage coefficients are I17 (*Arts, entertainment, and recreation*), I7 (*Retail trade*), and I18 (*Accommodation and food service*). Every additional one million dollars output produced by sector I5 (*Educational services*) will generate 67 jobs in the Illinois labor market.

By the same token, the breakdown of the highest per employment linkage coefficients of I13 (*Management of companies and enterprises*), I2 (*Mining*), and I11 (*Real estate and rental and leasing*) is shown in table 4.10. For every 100 jobs created/lost in I13 (*Management of companies and enterprises*), the economy in Illinois will eventually gain/lose a total of 760 jobs. Among them, 140 jobs will be in O17 (*Office and administrative support occupations*), 98 jobs will be in O1 (*Management occupations*), and 84 jobs will be in O16 (*Sales and related occupations*). In other words, the performance of I13 (*Management of companies and enterprises*) has quantitatively significant linkages to generate workforce externalities.

Table 4.8: Identification of Key Industries for Occupational Demand through Both Effects for Illinois State in 2005

	DIRECT EFFECT	INDIRECT EFFECT
	KEY INDUSTRY	KEY INDUSTRY
O1	I5 (*15.9%*)	I5 (*7.5%*)
O2	I10 (*25.2%*)	I5 (*9.6%*)
O3	I12 (*30.41%*)	I5 (*9.7%*)
O4	I5 (*41.7%*)	I5 (*7.9%*)
O5	I12 (*25.11%*)	I5 (*6.9%*)
O6	I16 (*37.7%*)	N.A.
O7	I12 (*63.1%*)	I5 (*13.1%*)
O8	I15 (*86.1%*)	N.A.
O9	I12 (*22.2%*), I9 (*18.2%*)	I5 (*9.5%*)
O10	I16 (*82.7%*)	N.A.
O11	I16 (*90.0%*)	N.A.
O12	I20 (*58.3%*)	N.A.
O13	I18 (*80.8%*)	I5 (*6.7%*)
O14	I14 (*35.9%*)	N.A.
O15	I16 (*33.9%*)	I5 (*4.5%*)
O16	I7 (*51.3%*)	I5 (*15.6%*)
O17	I10 (*12.6%*), I16 (*11.4%*)	I5 (*7.9%*)
O18	I1 (*79.4%*)	I5 (*20.3%*)
O19	I4 (*84.9%*)	N.A.
O20	I19 (*21.5%*), I5 (*19.5%*)	I5 (*7.9%*)
O21	I5 (*75.5%*)	I4 (*4.5%*)
O22	I8 (*35.4%*)	I5 (*8.2%*)

Table 4.9: Employment Linkage Coefficient and Output Linkage Coefficient in 2005 for Illinois State

	Per Employment Linkage Coefficient	Per Output Linkage Coefficient
I1	1.27	9.65
I2	2.77	2.45
I3	1.48	4.48
I4	1.36	11.97
I5	1.54	5.33
I6	1.50	6.44
I7	1.09	22.77
I8	1.27	9.63
I9	1.53	6.63
I10	1.27	5.49
I11	1.83	4.50
I12	1.41	7.48
I13	7.64	2.35
I14	1.18	13.18
I15	1.03	67.30
I16	1.19	16.64
I17	1.10	23.19
I18	1.14	22.44
I19	1.20	16.85
I20	1.20	3.40

Table 4.10: Breakdown of Per Employment Linkage Coefficients across 22-Occupation for Industries Showing Significant Linkage Effects in 2005 in Illinois

	I13	I2	I11
O1	0.98	0.25	0.26
O2	0.71	0.15	0.12
O3	0.45	0.06	0.03
O4	0.31	0.08	0.02
O5	0.11	0.05	0.01
O6	0.06	0.01	0.01
O7	0.37	0.05	0.03
O8	0.07	0.02	0.01
O9	0.29	0.08	0.02
O10	0.07	0.02	0.01
O11	0.02	0.00	0.00
O12	0.09	0.04	0.04
O13	0.18	0.06	0.05
O14	0.28	0.09	0.16
O15	0.16	0.06	0.02
O16	0.84	0.34	0.48
O17	1.40	0.35	0.30
O18	0.01	0.00	0.00
O19	0.33	0.40	0.05
O20	0.23	0.14	0.05
O21	0.39	0.25	0.07
O22	0.29	0.28	0.08
Sum	7.64	2.77	1.83

4.5 Concluding Remarks

The central purpose of this chapter has been to construct a general framework to trace the continuing transformations of a regional workforce, to understand and to guide that change through the linkages of occupations and industries. The analysis is a contribution to the larger body of research seeking to find measures to probe the linkages among industry-based policy, occupation-based targeting and regional workforce growth. The objective is not to forecast occupation shortages or surpluses or to provide estimates of occupational employment at each industry, but to estimate the growth signals from occupations and industries and to explore the fabric of occupation-industry linkage through sensitivity analysis. Outcomes and subsequent policy implications from the modeling have three major improvements over previous studies.

Firstly, labor market signaling focuses on the dynamic nature of skills demand and supply, and this chapter is geared toward examining forces that contribute to workforce changes over time. The results could be more informative, since identifying and calibrating the forces in action help understand how the structure of a local workforce is being shaped. Trends in the workforce through time can be detected to better guide policy-making decisions for both short-term and long-term bases. The proposed framework of linkage analysis between occupation and industry can be replicated in different time periods at different regional scales, and the occupation-by-industry matrices can be further disaggregated to detailed occupational and industrial levels.

Secondly, the monitoring of occupation-industry linkages can be used to detect labor market signals and further suggest appropriate schooling and training. This chapter demonstrates how forces underlying regional workforce growth can be decomposed and evaluated simultaneously. It differs from prior one-sided “industry-based” or “occupation-based” approaches and yields a number of rather acute insights into these rising and falling employment patterns. From the sensitivity and “key industry” analysis, a particular occupation/industry might be found to be strategically important in the development of an industry/occupation, and without it other occupations might not be created. There might be an upward or downward pressure on training and investment in some particular occupations. Therefore, the analysis guides manpower planners to revisit the social benefits from developing the strength of such an occupation and improve possible underinvestment in training. Moreover, the region-specific analysis of "per employment linkage coefficient" and "per employment output linkage coefficient" serves as a useful tool for development planners and policy makers to identify the priority of industries and justify state intervention to give incentives to these sectors for development purpose.

Thirdly, the model in this chapter takes a holistic approach to examine the composition of a regional workforce using a complete set of occupational groups and industrial sectors. It avoids the limitation of previous studies that rely heavily on pre-defined "creative occupations" without considering spatial heterogeneous distributions of occupation-based human capital and their functional roles. Particularly, the state-level analysis in this chapter helps clarify how divergent the growth force from an occupational group or an industrial sector could in comparison to the national level. The special features embedded in each state (such as existing occupation-mix or industry-mix, their endowment of human capital, policy environment for certain favored industrial sectors, etc.) could generate very different outcomes; these can be carefully incorporated into development strategy-making processes initiated by the policy makers. It appears that certain occupational groups or industrial sectors, not usually considered as "creative class" by Florida's standard, remain active in injecting vigor into the employment market. This may suggest that policy makers should not overlook the "non-creative" class and its functional role in an embedded economy. In some states, positive strong forces from O18 (*Farming, fishing, and forestry occupations*), I2 (*Mining*) and other supportive sectors do exist, and in this sense they deserve an equal recognition of their roles in promoting regional workforce growth. In other words, targeted occupations may not be limited to those held by highly educated professionals (Markusen, 2004).

Fourthly, it reveals possible cross-fertilization of occupational knowledge across industrial sectors for regional workforce growth. For example, the analysis revealed that while O18 (*Farming, fishing, and forestry occupations*) sends positive growing forces, sector I1 (*Agriculture, forestry, fishing, and hunting industry*) is experiencing negative growth. Obviously, farming, fishing, and forestry occupations are not the exclusive purview of agricultural, forestry, fishing, and hunting industry. The former provides a broad skill base to multiple industries, and the empirical analysis sends a clear message that the strength of skills/knowledge associated with agricultural work can be embedded in industries other than agricultural industries.

In common with other research work, analysis of regional workforce dynamics under industry-occupation linkage framework proposed in this chapter is not without limitations. First of all, the modeling structure simplifies the forces of employment growth into occupational and industrial functions. It does not explicitly consider scale effects and other less visible linkages. The numeric increase of occupational or industrial employment does not specify the quality of development. "Not all jobs are created equal; some pay a good deal more than others." (Florida *et al*., 2008) Secondly, categories of occupational groups and industrial sectors used in this chapter still remain at an aggregated level – i.e. a great deal of valuable, detailed information and dynamics, among more disaggregated occupational groups and industrial sectors, might have been overlooked in the findings. Also, the aggregated level of estimation results of the growth forces from the occupation side might not be particularly helpful in designing specific occupation training programs. This chapter, however, is demonstrative of the possibility of adopting an occupation-industry linkage analytical approach to study regional workforce dynamics. In future research, attention will be directed to these limitations.

REFERENCES

Adams, A. V., J. Middleton, and A. Ziderman (1992), Market-based manpower planning with labor market signals, *International Labor Review*, 131(3), 261-279.

Anselin, L. (2002), Under the hood: issues in the specification and interpretation of spatial regression models, *Agricultural Economics*, 27(3), 247-267.

Amos, O. M. Jr. (1989), An inquiry into the causes of increasing regional income inequality in the United State*s*, *Review of Regional Studies*, 19, 1-12.

Armstrong, H. (1995), Convergence among regions of the European Union, 1950-1990, *Papers in Regional Science*, 74(2), 143-152.

Barbour, E., and A. Markusen (2007), Regional occupational and industrial structure: does the one imply the other?, *International Regional Science Review*, 30(1), 72-90.

Barro, R. J. (1991), Economic growth in a cross-section of countries, *Quarterly Journal of Economics*, 106(2), 407-443.

Barro, R. J., and X. Sala-i-Martin (1992), Convergence, *Journal of Political Economy*, 100(2), 223 – 251.

Baumert, K., and M. Selman (2003), *Data Note: Heating and Cooling Degree Days*, World Resources Institute.

Bernard, A., and C. Jones (1996), Productivity and convergence across US states and industries, *Empirical Economics*, 21(1), 113-35.

Berndt, E. R., and D. O. Wood (1975), Technology, prices and the derived demand for energy, *The Review of Economics and Statistics*, 57(3), 259-268.

Berry, C. R., and E. L. Glaeser (2005), The divergence of human capital levels across cities, *Papers in Regional Science*, 84(3), 407 – 444.

Bivand, R., and R. Brunstad (2006), Regional growth in Western Europe: detecting spatial misspecification using the R environment, *Papers in Regional Science*, 85(2), 277-297.

Borts, G. H., and J. L. Stein (1964), *Economic Growth in a Free Market*, Columbia University Press, New York, NY.

Bredehoeft, J. D., and R. A. Young (1970), The temporal allocation of ground water – A simulation approach, *Water Resources Research*, 6(1), 3-21.

Brozovic, N., D. L. Sunding, and D. Zilberman (2007), Estimating business and residential water supply interruption losses from catastrophic events, *Water Resources Research*, 43, W08423.

Butler, J. J., X. Zhan, and V. A. Zlotnik (2007), Pumping-induced drawdown and stream depletion in a leaky aquifer system, *Ground Water*, 45(2), 178-186.

Carlino G., and L. Mills (1996), Convergence and the US States: a time series analysis, *Journal of Regional Science*, 36(4), 597-616.

Coe, D. T., and E. Helpman (1995), International R&D spillovers, *European Economic Review*, 39(5), 859-887.

Cohen, W., and D. Levinthal (1990), Absorptive capacity: a new perspective on learning and innovation, *Administrative Science Quarterly*, 35, 128-152.

Cohon, J. L. (1978), Multi-objective programming and planning, *Mathematics in Science and Engineering*, vol. 140, Academic Press Inc., London.

Coughlin, C. C., and T. B. Mandelbaum (1988), Why have state per-capita incomes diverged recently*? Review of the Federal Reserve Bank of St. Louis*, 70, 24-36.

Currid, E., and K. Stolarick (2010), The occupation-industry mismatch: new trajectories for regional cluster analysis and implications for economic development, *Urban Studies*, 47(2), 337-362.

Dalhuisen, J. M., R. J. G. M. Florax, H. L. F. M. de Groot, and P. Nijkamp (2003), Price and income elasticities of residential water demand: A meta-analysis, *Land Economics*, 79(2), 292-308.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, 6(2), 182-197.

De Mesnard, L. (1990), Biproportional method for analyzing interindustry dynamics: the case of France, *Economic Systems Research*, 2(3), 271-293.

De Mesnard, L. (1997), A biproportional filter to compare technical and allocation coefficient variations, *Journal of Regional Science*, 37(4), 541-564.

De Mesnard, L. (2000), Bicausative matrices to measure structural change: Are they a good tool?, *Annals of Regional Science*, 34(3), 421-449.

Dietzenbacher, E., J. A. Van Der Linden, and A. E. Steenge (1993), The regional extraction method: EC input-output comparisons, *Economic System Research*, 5(2), 185-206.

Domingues, E. P., E. A. Haddad, G. J. D. Hewings, and F. Perobelli (2002), Structural changes in the Brazilian interregional economic system, 1985-1997: Holistic matrix interpretation, *Australian Journal of Regional Studies*, 8(1), 21-44.

Duranton, G., D. Puga (2005), From sectoral to functional urban specialization, *Journal of Urban Economics*, 57(2), 343-370.

Easterlin, R. (1960a), Interregional differences in per capita income, population and total income, 1840-1950, in National Bureau of Economic Research (NBER) *Trends in the American Economy in the Nineteenth Century*. Princeton University Press, Princeton, NJ.

Easterlin, R. (1960b), Regional growth of income, in S. Kuznets, A. Miller, and R. Easterlin (Eds), *Population Redistribution and Economic Growth in the United States, 1870-1950*. American Philosophical Society, Philadelphia.

Elhorst, J. P. (2003), Specification and estimation of spatial panel data model, *International Regional Science Review*, 26(3), 244-268.

Ellerman, A. D., and B. K. Buchner (2007), The European Union Emission Trading Scheme: origins, allocation, and early results, *Review of Environmental Economics and Policy*, 1(1), 66-87.

Espey, M., J. Espey, and W. D. Shaw (1997), Price elasticity of residential demand for water: A meta-analysis, *Water Resources Research*, 33(6), 1369-1374.

Fan, C. C., and E. Casetti (1994), The spatial and temporal dynamics of US regional income inequality, 1950-1989, *Annals of Regional Science*, 28(2), 177-196.

Feinerman, E., and K. C. Knapp (1983), Benefits from groundwater management: magnitude, sensitivity, and distribution, *American Journal of Agricultural Economics*, 65(4), 703-710.

Feser, E. (2003), What regions do rather than make: a proposed set of knowledge-based occupation clusters, *Urban Studies*, 40(10), 1937 – 1958.

Fetter, C.W. (1988), *Applied Hydrogeology*. MacMillan, New York, NY.

Fingleton, B., and J. S. L. McCombie (1998), Increasing returns and economic growth: Some evidence for manufacturing from the European Union Regions, *Oxford Economic Papers*, 50(1), 89-105.

Fingleton, B. (1999), Estimates of time to economic convergence: an analysis of regions of the European Union, *International Regional Science Review,* 22(5), 5-35.

Fingleton, B. (2003), Externalities, economic geography and spatial econometrics: conceptual and modeling developments, *International Regional Science Review*, 26(2), 197-207.

Florida, R. (2002), *The Rise of Creative Class*. New York, Basic Books.

Florida, R., C. Mellander, and K. Stolarick (2008), Inside the black box of regional development – human capital, the creative class and tolerance, *Journal of Economic Geography*, 8(5), 615 – 649.

Glaeser, E., and A. Saiz (2004), The rise of skilled city, working paper, *Brookings – Wharton Papers on Urban Affairs*.

Glover, R. E., and G. G. Balmer (1954), River depletion resulting from pumping a well near a river, *Transactions of the American Geophysical Union*, 35(3), 468-470.

Gorelick, S. M. (1983), A review of distributed parameter groundwater management modeling methods, *Water Resources Research*, 19(2), 305-319.

Hantush, M. S. (1965), Wells near streams with semi-pervious beds, *Journal of Geophysical Research*, 70(12), 2829-2838.

Hayami, Y., and V. M. Ruttan (1985), *Agricultural Development: An International Perspective*, Baltimore and London: The Johns Hopkins University Press.

Hewitt, J. A., and W. M. Hanemann (1995), A discrete/continuous choice approach to residential water demand under block rate pricing, *Land Economics*, 71(2), 173-192.

Howe, C. W. (1982), The impact of price on residential water demand: Some new insights, *Water Resources Research*, 18(4), 713-716.

Hudson, E. A., and D. Jorgenson (1974), U.S. energy policy and economic growth, 1975-2000, *The Bell Journal of Economics and Management Science*, 5(2), 461-514.

Hunt, B. (1999), Unsteady stream depletion from ground water pumping, *Ground Water*, 37(1), 98-102.

Islam, N. (1995), Growth empirics: a panel data approach, *The Quarterly Journal of Economics*, 110(4), 1127-1170.

Jackson, R. W., P. Rogerson, D. Plane, and B. O. Huallachain (1990), A causative matrix approach to interpreting structural change, *Economic Systems Research*, 2(3), 261-270.

Jenkins, C. T. (1968), Techniques for computing rate and volume of stream depletion by wells, *Ground Water*, 6(2), 37-46.

Jenkins, M. W., J. R. Lund, and R. E. Howitt (2003), Using economic loss functions to value urban water scarcity in California, *Journal of American Water Works Association*, 95(2), 58-70.

Kalwij, I. M., and R. C. Peralta (2008), Non-adaptive and adaptive hybrid approaches for enhancing water quality management, *Journal of Hydrology*, 358(3-4):182-192

Kapoor, M., H. H. Kelejian, and I. R. Prucha (2007), Panel data models with spatially correlated error components, *Journal of Econometrics*, 140(1), 97-130.

Kelejian, H. H., and I. R. Prucha (1999), A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, 40(2), 509-533.

Keller, W. (2009), International trade, foreign direct investment, and technology spillovers, *NBER Working Paper Series* 15442, National Bureau of Economic Research, Inc.

King, K., C. Mellander, and K. Stolarick (2010), What you do, not who you work for – a comparison of the occupation industry structures of the United States, Canada and Sweden, *Electronic Working Paper Series*, Center of Excellence for Science and Innovation Studies, February 2010.

Koo, J. (2005), How to analyze the regional economy with occupation data, *Economic Development Quarterly*, 19(4), 356 – 372.

Krugman, P. (1983), *Targeted industrial policies: theory and evidence*, a paper prepared for the Conference on Industrial Change and Public Policy, sponsored by the Federal Reserve Bank of Kansas City, August 25-26, 1983, 46-49.

Kuznets, S. (1955), Economic growth and income inequality, *American Economic Review*, 45, 1-28.

Leser, C. E. V. (1954), The measurement of elasticities of demand, *Applied Statistics*, 3(2), 74-84.

Li, H., and K. E. Haynes (2011), Economic structure and regional disparity in China: beyond the Kuznets transition, *International Regional Science Review*, 34(2), 157-190.

Lipstein, B. (1968), Test marketing: a perturbation in the market place, *Management Science*, Series B, 14, 437-448.

Lopez-Bazo, E., E. V. Valcarce, A. J. Mora, and J. S. Caralt (1999), Regional economic dynamics and convergence in the European Union, *The Annals of Regional Science* 33(3): 343-370.

Lopez-Bazo, E., E. Vaya, and M. Artis (2004), Regional externalities and growth: evidence from European regions, *Journal of Regional Science*, 44(1), 43-73.

Lucas, R. (1988), On the mechanics of economic development, *Journal of Monetary Economics*, 22(1), 1 – 42.

Maddock, T. III (1974), The operation of a stream-aquifer system under stochastic demands, *Water Resource Research*, 10(1), 1 – 10.

Mankiw, N. G., D. Romer, and D. N. Weil (1992), A contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107(2), 407-437.

Manski, C. F. (1993), Identification of endogenous social effects: the reflection problem, *Review of Economic Studies*, 60 (3), 531-542.

Markusen, A. (2004), Targeting occupations in regional and community economic development, *Journal of the American Planning Association*, 70(3), 253-268.

Markusen, A., G. Schrock, and E. Barbour (2004), Making the city distinctive: a guide for planners and policymakers, Working Paper #159, Humphrey Institute of Public Affairs, University of Minnesota.

Markusen, A., and G. Schrock (2006), The distinctive city: divergent patterns in growth: hierarchy and specialization, *Urban Studies*, 43(8), 1301-1323.

Markusen, A. (2008), Human versus physical capital: government's role in regional development, in Jorge Martinez-Vazquez and Francois Vaillancourt, eds. *Public Policy for Regional Development*, Oxford, UK: Routledge.

Markandya, A., S. Pedroso-Galinato, and D. Streimikiene (2006), Energy intensity in transition economies: is there convergence towards the EU average? *Energy Economics*, 28(1), 121 – 145.

Mathur, V. K. (1999), Human capital-based strategy for regional economic development, *Economic Development Quarterly*, 13(3), 203-216.

Matlab Exchange, 2007 by Aravind Seshadri
http://www.mathworks.com/matlabcentral/fileexchange/10429

McGranahan, D., and T. Wojan (2007), Recasting the creative class to examine growth process in rural and urban counties, *Regional Studies*, 41(2), 197-216.

McHenry Hazard Mitigation Planning Committee (2010), McHenry County Natural Hazards Mitigation Plan,
http://www.state.il.us/iema/planning/Documents/Plan_McHenryCounty.pdf

Meller, P., and M. Marfan (1981), Small and large industry: employment generation, linkages and key sectors, *Economic Development and Cultural Change*, 29(2), 263-274.

Meyer, S. C., G. S. Roadcap, Y. F. Lin, and D. D. Walker (2009), *Kane County Water Resources Investigations: Simulation of Groundwater Flow in Kane County and Northeastern Illinois*, Illinois State Water Survey Contract Report 2009-07, Champaign, IL.

Miketa, A., and P. Mulder (2005), Energy productivity across developed and developing countries in 10 manufacturing sectors: patterns of growth and convergence, *Energy Economics*, 27(3), 429-453.

Mincer, J. (1974), *Schooling, Experiences, and Earnings*. New York: Columbia University Press.

Mueller, F. A. (1993), A management model for specification of groundwater withdrawal permits." *Water Resources Research*, 29(5), 1359 – 1368.

Mulder, P., and H. de Groot (2007), Sectoral energy and labor productivity convergence, *Environmental and Resource Economics*, 36(1), 85-112.

Mulder, P., and H. de Groot (2011), Energy-intensity developments for 19 OECD countries and 51 sectors, *CPB Discussion Paper*, 171, CPB Netherlands Bureau for Economic Policy Analysis.

Nieswiadomy, M. L. (1992), Estimating urban residential water demand: effects of price structure, conservation and education, *Water Resources Research*, 28(3), 609-615.

Park, W. (1995), International R&D spillovers and OECD economic growth, *Economic Inquiry*, 33(4), 571-591.

Patterson, M. G. (1996), What is energy efficiency?: Concepts, indicators and methodological issues, *Energy Policy*, 24(5), 377-390.

Piore, M. J., and C. F. Sabel (1984), *The Second Industrial Divide: Possibilities for Prosperity*. New York: Basic Books.

Price, L. (2005), *Voluntary Agreements for Energy Efficiency or GHG Emissions Reduction in Industry: An Assessment of Programs around the World*, Ernest Orland Lawrence Berkeley National Laboratory.

Rauch, J. E. (1993), Productivity gains from geographic concentration of human capital: evidence from the cities, *Journal of Urban Economics*, 34(3), 380 – 400.

Renwick, M. A., and R. D. Green (2000), Do residential water demand side management policies measure up? An analysis of eight California water agencies, *Journal of Environmental Economics and Management*, 40(1), 37-55.

Rey, S., and B. Montouri (1999), U.S. regional income convergence: A spatial econometric perspective, *Regional Studies*, 33(2), 143-156.

Rogerson, P., and D. Plane (1984), Modeling temporal changes in flow matrices, *Papers in Regional Science*, 54(1), 147-164.

Romer, P. (1986), Increasing returns and long-run growth, *Journal of Political Economy*, 94(5),

1002-1037.
Romer, P. (1987), Crazy explanations of the productivity slowdown, *NBER Maroeconomics Annual*, 2, 163 – 202.
Romer, P. (1990), Endogenous technological change, *Journal of Political Economy*, 98(5), S71 – S102.
Romer, P. (1994), The origins of endogenous growth, *Journal of Economic Perspectives*, 8(1), 3 – 22.
Sala-i-Martin, X. X. (1996), The classical approach to convergence analysis, *Economic Journal*, 106(437), 1019-1036.
Saxenian, A. (1994), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.
Scott, A. J. (2000), Economic geography: the great half-century, *Cambridge Journal of Economics*, 24(4), 1799-1817.
Solley, W. B., Pierce, R. R., and Perlman, H. A. (1998). "Estimated use of water in the United States in 1995." *U.S. Geological Survey Circular 1200*,
http://water.usgs.gov/watuse/pdf1995/pdf/circular1200.pdf.
Solow, R. M. (1956), A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70(1), 65-94.
Srinivas, N., and K. Deb (1994), Multiobjective optimization using nondominated sorting in genetic algorithms, *Evolutionary Computation*, 2(3), 221-248.
Stolarick, K., and R. Florida (2006), Creativity, connections and innovation: a study of linkages in the Montreal region, *Environmental and Planning A*, 38(10), 1799-1817.
Storper, M. (1997), *The Regional World: Territorial Development in a Global Economy*. New York: The Guilford Press.
Swan, T. W. (1956), Economic growth and capital accumulation, *Economic Record*, 32(2), 334-361.
Sweeney, S. H. (2004), Regional occupation employment projections: modeling supply constraints in the direct-requirements approach, *Journal of Regional Science*, 44(2), 263-288.
Theis, C. V. (1941), The effect of a well on the flow of a nearby stream, *Transactions American Geophysical Union*, 22(3), 734-738.
Thompson, W., and P. E. Thompson (1986), Accounting for municipal economic development,

Journal of Urban Affairs, 8 (2), 44 – 66.

Thompson, W., and P. R. Thompson (1987), National industries and local occupational strengths: the cross-hairs of targeting, *Urban Studies*, 24(6), 547-560.

Thompson, W. (1994), An occupational-functional approach to targeting industry, *Economic Development Review*, 12(2), 38 – 44.

Tiebout, C. M. (1956), Exports and regional economic growth, *Journal of Political Economy*, 64(2), 160-169.

Urpelainen, J. (2011), Export orientation and domestic electricity generation: effects on energy efficiency innovation in select sectors, *Energy Policy*, 39(9), 5638-5646.

Viscusi, A. E. (1995), Conflicting directives: water quality and appropriate water rights in the West, *William & Mary Environ. Law and Policy Review*, 20(5), 121-142.

Wan, J., E. Yang, and Y. F. Lin (2011), The Effect of groundwater allocation on economic welfare loss - a rationing policy study at a county scale under water scarcity, *Water Resources Management*, Under Review.

Ward, N. D., and P. Callander (2010), Distribution of stream depletion, *Journal of Hydrological Engineering - ASCE*, 136(3), 170-174.

Williamson, J. (1965), Regional inequality and the process of national development, *Economic Development & Cultural Change*, 13(4), 3-47.

Yeh, H., Y. Chang, and V. A. Zlotnik (2008), Stream depletion rate and volume from groundwater pumping in wedge-shape aquifers, *Journal of Hydrology*, 349(3-4), 501-511.

Young, R. A., and J. D. Bredehoeft (1972), Digital computer simulation for solving management problems of conjunctive groundwater and surface water system, *Water Resources Research*, 8(3), 533-556.