An Examination of Energy Intensity in the U.S. Manufacturing Sector

Gavin Pickenpaugh* and Peter Balash

*Pickenpaugh: National Energy Technology Laboratory, 3610 Collins Ferry Rd., Morgantown, WV 26507. Tel: (304) 285-0291. Email: gavin.pickenpaugh@netl.doe.gov. Balash: National Energy Technology Laboratory, 626 Cochrans Mill Rd. Pittsburgh, PA 15236. Tel: (412) 386-5753. Email: peter.balash@netl.doe.gov
1.0 INTRODUCTION

United States energy intensity, measured as primary energy consumption per dollar of real gross domestic product (GDP), fell approximately 45 percent from 1980 to 2009 (EIA, n.d.). Research concerning the driving forces of improvements in energy intensity dates back to the late 1970s, and a lot of the research attempts to tease out the contributions of efficiency and structural change (Myers and Nakamura, 1978). This paper is comprised of two major analysis sections, which cover the 1998, 2002, and 2006 periods for the U.S. manufacturing sector. The first analysis, conducted at both regional and national levels, decomposes movements in energy intensity into two factors: structural and efficiency change. The efficiency effect (also referred to in the literature as real intensity or technology effect) is the change in intensity resulting from efficiency changes within individual sectors. The structural effect (also referred to as change in product mix) is the change in energy intensity due to shifts in production between sectors. For instance, if efficiency of each individual sector remained the same, shifting production away from energy-intensive sectors results in lower energy intensity. Structural change can result from the migration of manufacturing to other countries, which may have competitive advantages, such as lower labor costs. For instance, in 2008, the average hourly compensation cost of manufacturing employees in China was $1.36, while the U.S. average was $32.26 (U.S. Department of Labor, 2011). Contributions of this section of this analysis to the literature are: 1. It uses a newer U.S. data set than has been used in previous papers (Manufacturing Energy Consumption Survey [MECS], which has data through 2006 and will likely be updated sometime in late 2012 with data from 2010); 2. It provides results at both national and regional levels, whereas much of the literature focuses on the national level.

The second analysis section uses regression models at both the national and regional levels to measure factors affecting energy intensity where the price of energy is the primary variable of interest; previous studies have found significant correlations between energy prices and intensity. This section examines the 1998-to-2006 time period (specifically 1998, 2002, and 2006), while other studies in the literature explore earlier and longer time periods.

This analysis adds to the literature by employing regression analysis as only a few of the studies to-date implement econometric methods. Unlike previous econometric literature on the subject, the present study uses 3-digit North American Industrial Classification System (NAICS) level data for manufacturing industries at the national and regional levels. Another contribution to the literature is that in addition to accounting for within-industry variation, between-industry variation is accounted for via random-effects and between models.

1 Several factors precluded this analysis from employing data from earlier time periods. The MECS conducted surveys prior to 1998, but used Standard Industrial Classification (SIC) during these earlier years. MECS data in 1998 employed both SIC and North American Industrial Classification System (NAICS) classifications; however, attempts to compare the two using SIC to NAICS mapping proved intractable. Additionally, from 1998 onward, NAICS was used in constructing Bureau of Economic Analysis (BEA) industry-level/state-level gross domestic product (GDP) numbers; Metcalf (2008) indicated that BEA advises against combining pre- and post-1997 state-level GDP data.
2.0 Literature Review

Two methods are generally used in the literature to tease out the components of changes in energy intensity: 1. Index decomposition analysis (IDA) and 2. Structural decomposition analysis (SDA). One advantage to IDA is that its data requirements are less stringent; SDA requires input output (I-O) tables, while IDA only requires sector-level data. Thus, the preferred method is largely dependent on the availability of data. Energy-related IDA studies often focus on total energy use or energy intensity. In the case of energy intensity, the components of changes in energy intensity are broken into two parts: 1. Efficiency change (also referred to as real intensity or technology effect) is the change in intensity resulting from efficiency changes within individual sectors. 2. Structural change (also referred to as change in product mix) is the change in energy intensity due to shifts in production between sectors. For instance, even if efficiency of any individual sector has not changed, shifting production away from energy-intensive sectors results in lower energy intensity. One example of structural change is the migration of manufacturing to other countries, which have advantages, such as lower labor costs. For example, in 2008, the average hourly compensation cost of manufacturing employees in China was $1.36, while the U.S. average was $32.26 (U.S. Department of Labor, 2008). Autor, et. al (2011) found that greater exposure to Chinese imports caused one-third of the reduction in U.S. manufacturing employment over the 1990-2007 period.

Studies estimating the causes of changes in energy intensity have been prevalent for several decades; Myers and Nakamura (1978) is cited in the literature as the earliest of these types of studies (Liu and Ang, 2007). Their study found the following: from 1967-74, energy intensity in the manufacturing sector was reduced one-third by structural change; from 1974-76, energy intensity was increased by one-third due to structural change. Implementing a decomposition method using U.S. four-digit SIC data, Samuel, et al. (1984) found that from 1975-80, of the 18.8 percent decline in energy intensity, over three-quarters was due to efficiency improvements, while a little under one-quarter was due to product mix changes (structural change). Energy use that excludes electricity fell in energy intensity by 20.9 percent, a little over three-fourths attributable to efficiency and almost one-fourth attributable to structural change. Electricity intensity fell 8.6 percent over the same time period, with efficiency driving almost all of the change.

Doblin (1988) decomposed U.S. energy intensity in the manufacturing sector. The study characterized the two decomposition components as structural change and technological change. From 1967 to 1974, energy intensity increased 0.4 percent, and structural changes had a greater effect than technology. From 1974 to 1980, energy intensity decreased over 2 percent per year; for this period, both structural change and technology change were found to have similar effects.

Liu and Ang (2007) conduct a literature review and discovered that the majority of U.S. studies found that both efficiency and structural change contribute to changes in energy intensity, with efficiency change having a larger impact in most studies. This is consistent with Huntington (1989), who indicated that studies examining time periods post-1960, on average, found about one-third of the decline in
energy intensity was due to structural change. However, not all studies concluded that efficiency change had a larger impact than structural change. The U.S. Department of Energy's (DOE) Office of Energy Efficiency and Renewable Energy (EERE) (2012) provides a spreadsheet with several intensity indexes that use the log Divisia method. In the manufacturing sector, there was a 31.8 percent decline in delivered energy intensity from 1986 to 2004. According to the EERE spreadsheet, structural change (23.8 percent decline) was responsible for about 70 percent of the decline, while efficiency change (10.5 percent improvement) was responsible for the remainder. Conversely, Choi and Ang (2012) examined the same data set and used a multiplicative Divisia index to estimate that about 65 percent of the decline was due to efficiency (22.3 percent improvement), while the rest was due to structural change (12.3 percent decline). One reason for the different findings using the same data is that Choi and Ang use a more aggregated sector scheme. It is unclear whether there are additional differences in calculation methods that result in contrasting conclusions. Another component of the EERE spreadsheet examines all industrial sectors (manufacturing and non-manufacturing); the 30.4 percent decline in aggregate intensity over the 1986-to-2004 time period was driven more by efficiency than by structural change. About 56 percent of the effect was due to structural change with the remaining 44 percent due to efficiency change. These comparisons show that the choice of sectors studied (i.e., all of manufacturing, the entire economy, non-manufacturing) can result in different conclusions.

Unander, et. al (1999) used an adaptive weighting Divisia energy IDA method for Organisation for Economic Co-operation and Development (OECD) countries, and for the U.S. manufacturing sector, they found that the efficiency effect explains more of the change in intensity than the structural effect during the 1981-94 time period. Mulder, et. al (2004) explores energy intensity at the country level and also at a sector level within countries by means of a Refined Divisia Index to measure the average annual macroeconomic energy and labor productivity growth rate as the sum of an efficiency effect and a structural effect. The country-level analysis found that virtually all of the growth of U.S. macroeconomic energy productivity was due to an efficiency effect. However, the sectoral analysis found that structural changes explained 29 percent of the movement in U.S. manufacturing energy intensity.

Boyd, et al. (1990) conducted an IDA Divisia decomposition analysis, and found that in the U.S. manufacturing sector, from 1974-81, structural change brought about a 1.2-percent-per-year decline in aggregate fuel intensity, while efficiency change was responsible for a 3.0-percent-per-year decline in aggregate intensity. They indicated that electricity intensity began to decrease in the 1970s and was driven mainly by increases in capacity utilization; the structural effect was small from 1970-81, averaging -0.5 percent per year, but became stronger from 1981 to 1985, averaging -2.8 percent per year.

Using the MECS, Energy Information Administration (1998) provided tables displaying the change in energy intensity as well as energy efficiency for U.S. manufacturing. From 1985 to 1994, aggregate energy intensity increased, while efficiency decreased in the U.S. manufacturing sector.

As apparent from the studies discussed hitherto, there are a variety of techniques used in IDA studies. Liu and Ang (2007) found that the logarithmic mean Divisia index (LMDI) methods are currently the preferred methods of the majority of researchers using IDA. One of the major benefits of the LMDI methods is the elimination of a residual term, which is often present in earlier studies. For instance,
Zarnikau (1999) used a Divisia approach to find that from 1981-88 there was a 16-percent change from efficiency and a 27.9-percent structural change; however, there was a large residual component of 22.2 percent. The elimination of the residual term in newer studies allows for a clearer interpretation of the effects ascribable to structural change and efficiency change. For example, if the residual component were zero, 43 percent \((16/(16+27.9))\) would be attributable to efficiency, with the rest ascribable to structural change.

There are other differences between studies, such as the level of sector aggregation. In the studies Liu and Ang (2007) examined, the number of subsectors ranged from 2 to 2,582, usually falling in the range of 5 to 25. They point out that changing the level of aggregation may produce different results. Different measures are also used in studies to calculate industry activity; gross output and value added have been the most used indicators of industrial activity.

IDA studies have been more common than SDA studies (Ang and Zhang, 2000). This is largely due to the more stringent data requirements of SDA. Additionally, SDA studies have been primarily used to investigate total energy use rather than energy intensity (Hoekstra et. al, 2003). As noted by Hoekstra et. al (2003), IDA and SDA energy studies have generally been conducted independently of one another. An exception is Weber (2009), who applied both SDA and IDA analysis and concluded that the structural effect was a bigger determinant than the efficiency effect in changes in U.S. energy intensity from 1997 to 2002.

Boyd, et al. (1990) and Weber (2008) concluded that the more disaggregated the sector scheme, the more energy use will appear to be driven by efficiency change. For instance, Boyd, et al. (1990) showed that when a two-sector disaggregation scheme was used, not as much of the sectoral shift will be captured as in a more disaggregated scheme. The reasoning is that in a relatively aggregated scheme, changes in structure at more disaggregated levels are not accounted for; thus, relevant changes in intensity would incorrectly be attributed to efficiency change rather than structural change.

There have been studies that go beyond the standard IDA and SDA analysis of energy intensity. Choi, et al. (1995) proposed a method to decompose energy intensity into structural change, interfuel substitution, and efficiency change. Using a decomposition procedure, Wing (2008) found that from 1958 to 1979, U.S. energy intensity was driven more by structural change than efficiency change; however, from 1980 to 2000, efficiency change was responsible for almost two-thirds of the decline in energy intensity. In addition, the study used an econometric approach, which breaks out efficiency change into three major areas: 1. Substitution associated with variable inputs, 2. Disembodied technical progress, and 3. Quasi-fixed inputs. The study found that changes in the composition of quasi-fixed inputs were the largest driver of real efficiency change over both the 1958-79 and the 1980-2000 periods. Wing and Eckaus (2004) used a decomposition method, finding that changes in efficiency were the major driver of U.S. energy intensity after 1980. Via an econometric approach, they found that changes in prices had an insignificant effect on efficiency prior to 1974, but a significant effect afterwards, with the largest effect occurring during the 1974-86 time period. The study also concluded that quasi-fixed inputs caused a decline in efficiency before 1986, but lead to improved efficiency thereafter.
Metcalf (2008) used IDA at the U.S. national and state level and subsequently used fixed-effects models to estimate the contribution of changes in economic and climate variables to energy intensity, as well as structural change and efficiency change. Using national level data, the study estimated that changes in efficiency from 1970 to 2003 caused around 75 percent of the decline in U.S. energy intensity. The structural effect increased post-1990, but efficiency changes remained the key driver in bringing about intensity changes. The state data analysis found changes in efficiency from 1960 to 2001 caused about 64 percent of the decline in intensity. The econometric analysis concluded that higher per capita income and energy prices were associated with improved energy efficiency (as well as energy intensity), but their effect on structural change was not significant, implying that price and income changes influence intensity via the efficiency effect, rather than structural change.

Bernstein, et. al (2003) did not conduct a decomposition analysis; instead, the study employed fixed-effects regression models to estimate which factors affected energy intensity in the U.S. from 1977 to 1999. The study determined that in the industrial sector, increases in energy prices were associated with lower energy intensities. Using dummy variables for individual energy intensive sectors, the study also concluded that growth in energy-intensive subsectors results in higher energy intensity.

Alghandoor, et al. (2008) implemented a regression analysis to decompose energy intensity in the U.S. manufacturing sector, using 1977-98 SIC data. The regression analysis estimated that structural change contributed to decreases of 28 percent of the total decline in aggregate fuel intensity and 41 percent of the decline in aggregate electricity intensity over the period. Results were found to be similar to IDA results.
3.0 Data

Data cover the 1998, 2002, and 2006 time periods. Descriptions of the variables used in the paper are provided in Table 1.

Table 1. Variable Descriptions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Units</th>
<th>Level of Detail</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Real Gross Domestic Product (GDP)</td>
<td>Million Chained 2005$</td>
<td>State</td>
<td>BEA: Regional Data, GDP &amp; Personal Income</td>
</tr>
<tr>
<td>Cons</td>
<td>First use of energy for all purposes (fuel and nonfuel)</td>
<td>Trillion Btu</td>
<td>Region</td>
<td>MECS, Table 1.2</td>
</tr>
<tr>
<td>EI</td>
<td>Energy Intensity (EI), equal to Cons/GDP</td>
<td>Trillion Btu/Million Chained 2005$</td>
<td>Region</td>
<td>Calculation from above variables</td>
</tr>
<tr>
<td>Price</td>
<td>Average price of purchased energy sources</td>
<td>Dollars per MMbtu (2011$)</td>
<td>Region</td>
<td>MECS, Table 7.2; BEA: NIPA Table 1.1.9</td>
</tr>
<tr>
<td>Partic</td>
<td>Participation Rate in Energy Management Activities</td>
<td>Percent</td>
<td>National</td>
<td>MECS, Table 8.1</td>
</tr>
<tr>
<td>Estab</td>
<td>Number of establishments</td>
<td>Number of establishments</td>
<td>National</td>
<td>MECS, Table 9.1</td>
</tr>
<tr>
<td>Avgflsp</td>
<td>Average floorspace per establishment</td>
<td>sqft/establishment</td>
<td>National</td>
<td>MECS, Table 9.1</td>
</tr>
<tr>
<td>Empl</td>
<td>Total full-time and part-time employment</td>
<td>Number of employees</td>
<td>State</td>
<td>BEA: Regional Data, GDP &amp; Personal Income, Table SA25N</td>
</tr>
<tr>
<td>Wage</td>
<td>Wage and salary disbursements</td>
<td>Thousand Real 2011$</td>
<td>State</td>
<td>BEA: Regional Data, GDP &amp; Personal Income, Table SA07N; BEA: NIPA Table 1.1.9</td>
</tr>
<tr>
<td>Avgwage</td>
<td>Average wage, equal to Wage/Empl</td>
<td>Thousand Real 2011$/employee</td>
<td>State</td>
<td>Calculation from above variables</td>
</tr>
<tr>
<td>Avgempl</td>
<td>Average employees, equal to Employment/Estab</td>
<td>employees/estab</td>
<td>National</td>
<td>Calculation from above variables</td>
</tr>
</tbody>
</table>
The industry breakout is summarized in Table 2. Notice that the top three rows are aggregates of two industries (e.g., 311_312 combines NAICS 311 and NAICS 312).

Table 2. Industry Scheme

<table>
<thead>
<tr>
<th>NAICS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>311_312</td>
<td>Food and Beverage mfg</td>
</tr>
<tr>
<td>313_314</td>
<td>Textile and Textile Product mills</td>
</tr>
<tr>
<td>315_316</td>
<td>Apparel and Leather and Allied Product mfg</td>
</tr>
<tr>
<td>321</td>
<td>Wood product mfg</td>
</tr>
<tr>
<td>322</td>
<td>Paper mfg</td>
</tr>
<tr>
<td>323</td>
<td>Printing and related support activities</td>
</tr>
<tr>
<td>324</td>
<td>Petroleum and coal products mfg</td>
</tr>
<tr>
<td>325</td>
<td>Chemical mfg</td>
</tr>
<tr>
<td>326</td>
<td>Plastics and rubber products mfg</td>
</tr>
<tr>
<td>327</td>
<td>Nonmetallic mineral product mfg</td>
</tr>
<tr>
<td>331</td>
<td>Primary metal mfg</td>
</tr>
<tr>
<td>332</td>
<td>Fabricated metal product mfg</td>
</tr>
<tr>
<td>334</td>
<td>Computer and electronic product mfg</td>
</tr>
<tr>
<td>335</td>
<td>Electrical equipment, appliance, and component mfg</td>
</tr>
<tr>
<td>336</td>
<td>Motor vehicle, body, trailer, and parts mfg</td>
</tr>
<tr>
<td>336</td>
<td>Other transportation equipment mfg</td>
</tr>
<tr>
<td>337</td>
<td>Furniture and related product mfg</td>
</tr>
<tr>
<td>339</td>
<td>Miscellaneous mfg</td>
</tr>
</tbody>
</table>
3.1 Regional Breakout

Figure 1 displays the regional boundaries.

Figure 1. U.S. Regions
3.2 Summary Statistics

Table 3 provides summary statistics of the national level variables. Three different standard deviation measures are used: 1. Overall, 2. Between, and 3. Within. The Between measure only captures variation across NAICS codes, while the Within measure only captures variation within NAICS codes. The Overall measure encompasses both within and between variation. Each variable demonstrates substantially more between variation than within variation.

Table 4 displays the correlations between the variables (levels). Of particular interest, Price and energy intensity (EI) have a -0.68 correlation, indicating higher energy prices are associated with lower energy intensities. Another interesting relationship is the 0.57 correlation between intensity and participation. Industries with higher participation rates in energy efficiency programs tend to have higher energy intensities than industries with lower participation; it is likely more attractive for industries that are less efficient to enroll in such programs since they may have more to gain from improvements in efficiency than industries that have relatively low-energy intensities.

Table 5 displays the correlations between the first differences of the national-level variables. The correlations of the levels (Table 4) and first differences (Table 5) tend to differ. Of particular interest is the weak (-0.04), although still negative, correlation between the first differences of EI and Price. Combined with the findings from Table 4, this implies that while cross-section correlation between Price and Intensity is high, the correlation of changes within industries over time is much smaller for the 1998, 2002, and 2006 periods. Another major difference between the first difference and level correlations is the negative correlation between the first differences EI and Avgwage of -0.61, while the levels of the two variables have a positive correlation (0.52). Also of note is the weak (positive) correlation between the first differences of EI and Partic, suggesting within-industry increases in participation in efficiency programs may not have a large association with within-industry changes in energy intensity.

Tables 6, 9, 12, and 15 provide the regional summary statistics. As in the national level data, between variation tends to exceed within variation. Tables 7, 10, 13, and 16 provide the regional correlation matrices of the levels of the variables. Correlations between Price and EI are similar to the national figures, ranging from -0.56 to -0.63. (Note: Several of the national level variables are not available at the regional level.) Tables 8, 11, 14, and 17 provide the regional correlation matrices of the first differences of the variables. One finding of note is the small and positive (in three of the regions) correlation between the first difference of EI and Price; the correlations range from -0.19 to +0.36.
Table 3. National Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Observations^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>80,081.2</td>
<td>57,846.8</td>
<td>15,670.0</td>
<td>223,815.0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>55,189.1</td>
<td>20,380.1</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>1,251.2</td>
<td>1,993.9</td>
<td>17</td>
<td>7,320.0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>2,025.1</td>
<td>171.2</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Cons</td>
<td>1,251.2</td>
<td>1,993.9</td>
<td>17</td>
<td>7,320.0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>2,025.1</td>
<td>171.2</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>EI</td>
<td>0.0154</td>
<td>0.0207</td>
<td>0.0006</td>
<td>0.0895</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>0.0207</td>
<td>0.0039</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Price</td>
<td>11.3</td>
<td>4.4</td>
<td>3.2</td>
<td>19.9</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>3.9</td>
<td>2.1</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Partic</td>
<td>0.4300</td>
<td>0.1222</td>
<td>0.1</td>
<td>0.7</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>0.1124</td>
<td>0.0527</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Estab</td>
<td>11,523.3</td>
<td>8,264.5</td>
<td>1,756.0</td>
<td>40,743.0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>8,306.8</td>
<td>1,379.7</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Avgflsp</td>
<td>99,371.4</td>
<td>65,338.6</td>
<td>20,605.7</td>
<td>356,694.8</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>58,671.9</td>
<td>30,929.7</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Empl</td>
<td>898,940.7</td>
<td>521,245.5</td>
<td>115,000.0</td>
<td>2,090,900.0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>521,619.7</td>
<td>99,384.0</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Wage</td>
<td>47,200,000</td>
<td>34,500,000</td>
<td>9,449,209</td>
<td>131,000,000</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>35,000,000</td>
<td>2,993,588</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Avgwage</td>
<td>51.4</td>
<td>15.8</td>
<td>26.2</td>
<td>92.6</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>15.9</td>
<td>3.0</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Avgempl</td>
<td>93.9</td>
<td>53.7</td>
<td>34.7</td>
<td>252.4</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>54.3</td>
<td>6.9</td>
<td></td>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

^2 The number of between observations is the average number of observations in each time period (e.g., 1998 GDP has 18 observations). The number of within observations is the average number of observations for each NAICS (e.g., NAICS 311 has three observations—one observation for each of 1998, 2002, and 2006).
Table 4. National Correlation Matrix (Levels)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Price</th>
<th>Partic</th>
<th>Avgflsp</th>
<th>Avgwage</th>
<th>Avgempl</th>
<th>Cons</th>
<th>Estab</th>
<th>Wage</th>
<th>Empl</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>0.01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.68</td>
<td>0.10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partic</td>
<td>0.57</td>
<td>0.37</td>
<td>-0.37</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgflsp</td>
<td>0.19</td>
<td>0.21</td>
<td>-0.49</td>
<td>0.50</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>0.52</td>
<td>0.57</td>
<td>-0.23</td>
<td>0.54</td>
<td>0.28</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgempl</td>
<td>0.01</td>
<td>0.46</td>
<td>-0.18</td>
<td>0.40</td>
<td>0.72</td>
<td>0.48</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.87</td>
<td>0.33</td>
<td>-0.55</td>
<td>0.58</td>
<td>0.16</td>
<td>0.65</td>
<td>0.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estab</td>
<td>-0.47</td>
<td>0.18</td>
<td>0.45</td>
<td>-0.48</td>
<td>-0.49</td>
<td>-0.30</td>
<td>-0.42</td>
<td>-0.36</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.33</td>
<td>0.74</td>
<td>0.27</td>
<td>0.02</td>
<td>0.20</td>
<td>0.45</td>
<td>0.64</td>
<td>-0.12</td>
<td>0.29</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.49</td>
<td>0.67</td>
<td>0.36</td>
<td>-0.11</td>
<td>0.07</td>
<td>0.12</td>
<td>0.45</td>
<td>-0.29</td>
<td>0.54</td>
<td>0.91</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5. National Correlation Matrix (First Differences)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Price</th>
<th>Partic</th>
<th>Avgflsp</th>
<th>Avgwage</th>
<th>Avgempl</th>
<th>Cons</th>
<th>Estab</th>
<th>Wage</th>
<th>Empl</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.28</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.04</td>
<td>0.26</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partic</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.18</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgflsp</td>
<td>-0.03</td>
<td>-0.05</td>
<td>-0.25</td>
<td>-0.24</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>-0.61</td>
<td>0.68</td>
<td>0.10</td>
<td>0.10</td>
<td>-0.08</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgempl</td>
<td>0.15</td>
<td>-0.04</td>
<td>0.08</td>
<td>0.31</td>
<td>0.12</td>
<td>0.02</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.28</td>
<td>0.04</td>
<td>-0.11</td>
<td>0.06</td>
<td>-0.11</td>
<td>0.02</td>
<td>0.13</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estab</td>
<td>-0.12</td>
<td>0.15</td>
<td>0.22</td>
<td>-0.12</td>
<td>-0.12</td>
<td>0.03</td>
<td>-0.41</td>
<td>-0.07</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.09</td>
<td>0.09</td>
<td>0.31</td>
<td>0.25</td>
<td>-0.08</td>
<td>-0.01</td>
<td>0.35</td>
<td>0.05</td>
<td>0.45</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.10</td>
<td>-0.13</td>
<td>0.28</td>
<td>0.07</td>
<td>-0.10</td>
<td>-0.24</td>
<td>0.19</td>
<td>-0.04</td>
<td>0.57</td>
<td>0.89</td>
<td>1</td>
</tr>
</tbody>
</table>
### Table 6. West Region Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP</strong></td>
<td>overall</td>
<td>14,831</td>
<td>16,107</td>
<td>1,821</td>
<td>102,300</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>8,584</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>13,893</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>overall</td>
<td>142.3</td>
<td>266.7</td>
<td>1.0</td>
<td>1,215.0</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>267.1</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>21.6</td>
<td></td>
<td></td>
<td>2.89*</td>
</tr>
<tr>
<td><strong>EI</strong></td>
<td>overall</td>
<td>0.0106</td>
<td>0.0149</td>
<td>0.0002</td>
<td>0.0563</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>0.0148</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>0.0027</td>
<td></td>
<td></td>
<td>2.89*</td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>overall</td>
<td>14.2</td>
<td>7.6</td>
<td>2.3</td>
<td>32.1</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>6.5</td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>4.0</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>Empl</strong></td>
<td>overall</td>
<td>169,490</td>
<td>140,247</td>
<td>24,176</td>
<td>673,987</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>141,408</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>20,679</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>Wage</strong></td>
<td>overall</td>
<td>9,431,086</td>
<td>11,800,000</td>
<td>1,244,959</td>
<td>53,800,000</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>12,100,000</td>
<td></td>
<td>499,310</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>499,310</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>Avgwage</strong></td>
<td>overall</td>
<td>51.4</td>
<td>20.0</td>
<td>22.9</td>
<td>111.9</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>20.0</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>4.1</td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

*Average number of time periods (less than 3, due to missing data)
Table 7. West Region Correlation Matrix (Levels)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.73</td>
<td>0.19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.60</td>
<td>0.00</td>
<td>-0.47</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.40</td>
<td>0.64</td>
<td>-0.26</td>
<td>0.14</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.30</td>
<td>0.74</td>
<td>-0.17</td>
<td>0.09</td>
<td>0.92</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>0.32</td>
<td>0.66</td>
<td>0.59</td>
<td>-0.24</td>
<td>0.24</td>
<td>0.49</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 8. West Region Correlation Matrix (First Differences)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.17</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>-0.07</td>
<td>0.13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>0.08</td>
<td>0.10</td>
<td>0.04</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.06</td>
<td>-0.57</td>
<td>-0.03</td>
<td>0.13</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.10</td>
<td>0.10</td>
<td>0.01</td>
<td>0.19</td>
<td>0.70</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>-0.45</td>
<td>0.63</td>
<td>0.20</td>
<td>-0.12</td>
<td>-0.35</td>
<td>0.14</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 9. South Region Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP</strong></td>
<td>overall</td>
<td>27,826</td>
<td>20,920</td>
<td>5,098</td>
<td>84,682</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>19,991</td>
<td>7,285</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>overall</td>
<td>734.4</td>
<td>1,426.9</td>
<td>12.0</td>
<td>5,417.0</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>1,449.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td>115.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EI</strong></td>
<td>overall</td>
<td>0.0210</td>
<td>0.0310</td>
<td>0.0006</td>
<td>0.1340</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td></td>
<td>0.0305</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td>0.0078</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>overall</td>
<td>10.6</td>
<td>4.2</td>
<td>1.7</td>
<td>20.7</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td></td>
<td>3.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td>2.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Empl</strong></td>
<td>overall</td>
<td>299,635</td>
<td>150,263</td>
<td>53,298</td>
<td>647,795</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td></td>
<td>147,983</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td>38,811</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Wage</strong></td>
<td>overall</td>
<td>14,300,000</td>
<td>8,350,559</td>
<td>2,925,753</td>
<td>32,800,000</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td></td>
<td>8,427,589</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td>1,178,025</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Avgwage</strong></td>
<td>overall</td>
<td>48.7</td>
<td>15.8</td>
<td>24.8</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td>between</td>
<td></td>
<td>15.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td></td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 10. South Region Correlation Matrix (Levels)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>0.21</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.88</td>
<td>0.47</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.63</td>
<td>0.00</td>
<td>-0.48</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.46</td>
<td>0.54</td>
<td>-0.27</td>
<td>0.29</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.24</td>
<td>0.62</td>
<td>-0.02</td>
<td>0.31</td>
<td>0.85</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>0.61</td>
<td>0.44</td>
<td>0.66</td>
<td>-0.11</td>
<td>-0.20</td>
<td>0.27</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 11. South Region Correlation Matrix (First Differences)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.49</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.60</td>
<td>-0.24</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.19</td>
<td>0.37</td>
<td>-0.21</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.17</td>
<td>0.10</td>
<td>-0.04</td>
<td>0.20</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.12</td>
<td>0.19</td>
<td>-0.00</td>
<td>0.15</td>
<td>0.92</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>-0.56</td>
<td>0.70</td>
<td>-0.07</td>
<td>-0.11</td>
<td>-0.08</td>
<td>0.03</td>
<td>1</td>
</tr>
<tr>
<td>Variable</td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Min</td>
<td>Max</td>
<td>Observations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>-----------</td>
<td>--------</td>
<td>---------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>24,072</td>
<td>20,627</td>
<td>1,462</td>
<td>86,318</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>20,720</td>
<td>1,571</td>
<td>80,307</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>3,514</td>
<td>14,366</td>
<td>36,896</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>281.0</td>
<td>339.4</td>
<td>1.0</td>
<td>1,284.0</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>339.7</td>
<td>2.7</td>
<td>1,090.0</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>51.5</td>
<td>92.0</td>
<td>475.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EI</td>
<td>0.0137</td>
<td>0.0195</td>
<td>0.0006</td>
<td>0.0844</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>0.0194</td>
<td>0.0011</td>
<td>0.0751</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>0.0030</td>
<td>0.0073</td>
<td>0.0230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>10.3</td>
<td>3.6</td>
<td>4.0</td>
<td>17.0</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>3.3</td>
<td>5.3</td>
<td>14.9</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>1.5</td>
<td>7.2</td>
<td>14.0</td>
<td>2.94*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>278,497</td>
<td>218,098</td>
<td>21,689</td>
<td>929,895</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>219,731</td>
<td>24,003</td>
<td>831,185</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>33,272</td>
<td>178,111</td>
<td>377,208</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>14,800,000</td>
<td>13,300,000</td>
<td>1,064,913</td>
<td>60,100,000</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>13,500,000</td>
<td>1,330,856</td>
<td>55,500,000</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>1,551,774</td>
<td>9,122,347</td>
<td>19,800,000</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>50.6</td>
<td>12.6</td>
<td>27.4</td>
<td>79.0</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>between</td>
<td>12.7</td>
<td>29.3</td>
<td>73.4</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>2.1</td>
<td>45.2</td>
<td>57.7</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Average number of time periods (less than 3, due to missing data)
Table 13. Midwest Region Correlation Matrix (Levels)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.21</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.85</td>
<td>0.15</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.59</td>
<td>-0.05</td>
<td>-0.66</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.41</td>
<td>0.87</td>
<td>-0.12</td>
<td>0.15</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.34</td>
<td>0.91</td>
<td>-0.05</td>
<td>0.10</td>
<td>0.96</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>0.47</td>
<td>0.42</td>
<td>0.58</td>
<td>-0.26</td>
<td>0.10</td>
<td>0.30</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 14. Midwest Region Correlation Matrix (First Differences)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.28</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.27</td>
<td>0.30</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>0.36</td>
<td>0.09</td>
<td>0.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>0.10</td>
<td>0.22</td>
<td>0.13</td>
<td>0.28</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>0.09</td>
<td>0.30</td>
<td>0.15</td>
<td>0.21</td>
<td>0.95</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>-0.14</td>
<td>0.31</td>
<td>0.27</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.17</td>
<td>1</td>
</tr>
<tr>
<td>Variable</td>
<td>Overall</td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Min</td>
<td>Max</td>
<td>Observations</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-------</td>
<td>-----------</td>
<td>------</td>
<td>-------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>overall</td>
<td>13,353</td>
<td>11,411</td>
<td>2,619</td>
<td>55,226</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>between</td>
<td>11,238</td>
<td>3,418</td>
<td>50,717</td>
<td>18,111</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>within</td>
<td>2,950</td>
<td>1,919</td>
<td>55,226</td>
<td>2,611</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>overall</td>
<td>109.1</td>
<td>143.8</td>
<td>2.0</td>
<td>612.0</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>between</td>
<td>142.9</td>
<td>7.1</td>
<td>555.7</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>within</td>
<td>23.9</td>
<td>41.1</td>
<td>194.6</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>EI</td>
<td>overall</td>
<td>0.0113</td>
<td>0.0182</td>
<td>0.0005</td>
<td>0.0919</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>EI</td>
<td>between</td>
<td>0.0177</td>
<td>0.0014</td>
<td>0.0721</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>EI</td>
<td>within</td>
<td>0.0044</td>
<td>-0.0048</td>
<td>0.0311</td>
<td>2.89*</td>
<td>2.89*</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>overall</td>
<td>14.3</td>
<td>5.6</td>
<td>3.9</td>
<td>26.5</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>between</td>
<td>4.8</td>
<td>5.2</td>
<td>20.7</td>
<td>17</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>within</td>
<td>3.0</td>
<td>7.8</td>
<td>25.1</td>
<td>2.94*</td>
<td>2.94*</td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>overall</td>
<td>151,665</td>
<td>86,123</td>
<td>14,911</td>
<td>366,203</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>between</td>
<td>85,025</td>
<td>16,230</td>
<td>316,609</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>within</td>
<td>21,463</td>
<td>104,391</td>
<td>213,389</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>overall</td>
<td>8,512,942</td>
<td>6,257,498</td>
<td>1,147,577</td>
<td>25,000,000</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>between</td>
<td>6,327,872</td>
<td>1,226,903</td>
<td>24,000,000</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>within</td>
<td>790,785</td>
<td>6,670,536</td>
<td>11,000,000</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>overall</td>
<td>53.9</td>
<td>14.6</td>
<td>31.0</td>
<td>90.6</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>between</td>
<td>14.7</td>
<td>32.8</td>
<td>84.6</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>within</td>
<td>2.4</td>
<td>45.6</td>
<td>60.1</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

*Average number of time periods (less than 3, due to missing data)
Table 16. Northeast Region Correlation Matrix (Levels)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.25</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.92</td>
<td>0.02</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-0.56</td>
<td>0.02</td>
<td>-0.66</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>-0.52</td>
<td>0.61</td>
<td>-0.37</td>
<td>0.31</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>-0.41</td>
<td>0.73</td>
<td>-0.23</td>
<td>0.26</td>
<td>0.92</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>0.31</td>
<td>0.60</td>
<td>0.47</td>
<td>-0.10</td>
<td>0.26</td>
<td>0.57</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 17. Northeast Region Correlation Matrix (First Differences)

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>GDP</th>
<th>Cons</th>
<th>Price</th>
<th>Empl</th>
<th>Wage</th>
<th>Avgwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-0.29</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>0.54</td>
<td>-0.13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>0.15</td>
<td>0.14</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empl</td>
<td>0.21</td>
<td>-0.20</td>
<td>0.01</td>
<td>0.14</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>0.18</td>
<td>0.07</td>
<td>0.18</td>
<td>0.04</td>
<td>0.81</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Avgwage</td>
<td>0.03</td>
<td>0.53</td>
<td>0.17</td>
<td>-0.05</td>
<td>-0.41</td>
<td>0.09</td>
<td>1</td>
</tr>
</tbody>
</table>
4.0 Decomposition Analysis

This section provides the methodology and results for the decomposition component of the analysis. One contribution of this analysis to the literature is it uses a newer U.S. data set than has been used in previous papers (MECS, which has data through 2006). Another contribution is that a U.S. regional analysis was conducted; whereas, much of the literature used national-level data.

4.1 National Level Methodology

Similar to most of the studies in the literature review, this report employs a Divisia IDA model to tease out the components of change in energy intensity: structural and efficiency change. It is important to keep in mind the data in this study is only manufacturing industries; for instance, this implies that shifts between various service industries are not accounted for. Using the Multiplicative Log-Divisia Index, the relevant formulas for decomposing the movements in energy intensity between two time periods are the following (Choi and Ang, 2012):

\[
\frac{E_{i_t}}{E_{i_{98}}} = \frac{R_i}{R_{98}} \times \frac{S_i}{S_{98}}
\]

\[
\frac{R_i}{R_{98}} = \exp\left(\sum_{i=1}^{n} w_i \times (\ln(E_{i,t}) - \ln(E_{i,98}))\right)
\]

\[
\frac{S_i}{S_{98}} = \exp\left(\sum_{i=1}^{n} w_i \times (\ln(S_{i,t}) - \ln(S_{i,98}))\right)
\]

\[
w_i = \frac{(E_{i,t} \times S_{i,t}) - (E_{i,98} \times S_{i,98}) / ((\ln(E_{i,t}) \times S_{i,t}) - (\ln(E_{i,98}) \times S_{i,98}))}{(E_{i,t} - E_{i,98}) / ((\ln(E_{i,t}) - (\ln(E_{i,98}))}
\]

EI: Energy intensity index

\(R_i/R_{98}\): Real Efficiency index (1998 is the base year)

\(S_i/S_{98}\): Structural index (1998 is the base year)

\(S_{i,t}\): Production share of sector i at time t

\(w_i\): Weight for industry i

t: time period t

98: time period equals 1998

exp: exponential

Ln: natural logarithm
4.2 Regional Level Methodology

In a manner similar to the National Level Decomposition, decompositions are carried out for each of the four regions.

Due to lack of data, industries 313_314 and 315_316 are excluded from the regional analysis (except the South region, which had sufficient data for industry 313_314). Additionally, the West and Northeast regions do not have sufficient data for industry 331, while the Northeast region lacks sufficient data for industry 333.

4.3 National Level Decomposition Results

Decomposition results at the national level are displayed in Figure 2. Aggregate energy intensity declined 28.8 percent from 1998 to 2006. This was primarily driven by real efficiency, which improved 24.8 percent over the same period. Structural changes declined 5.3 percent over the same period. This implies real efficiency changes were responsible for 82 percent of the decline in aggregate intensity, while structural change was responsible for the remaining 18 percent. (Note: Declines in the real efficiency index mean that the individual industries became more efficient, as opposed to less efficient. This may not be readily apparent from looking at the graph.)

Figure 2. National Decomposition Indexes
4.4 Regional Level Decomposition Results

Conducting the decomposition analysis at regional levels allows a way to further examine the national findings. Figures 3, 4, 5, and 6 summarize the regional decomposition results. Results vary significantly by region. In stark contrast to the national model, the Northeast and Midwest regions’ declines in aggregate intensity were driven by structural change, rather than efficiency change; conversely, the South and West regions’ declines in intensity were driven by real efficiency change. *(Note: Declines in the real efficiency index mean that the individual industries became more efficient, as opposed to less efficient.)*

More specifically, energy intensity in the West Region declined 45.2 percent from 1998 to 2006. The decomposition results imply 65 percent of this decline is ascribable to real efficiency improvements. The entire 31.3 percent decline in South energy intensity is attributable to improvements in real efficiency. Conversely, 79 percent of the Midwest intensity decline of 18.3 percent is attributable to structural change, with only 21 percent ascribable to efficiency improvements. In the Northeast, 81 percent of the 32 percent decline in energy intensity is attributable to structural change, with the remaining 19 percent attributable to efficiency improvements.

**Figure 3. West Region Decomposition Indexes**
Figure 4. South Region Decomposition Indexes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Efficiency</td>
<td>98.6</td>
<td>67.6</td>
<td></td>
</tr>
<tr>
<td>Structural</td>
<td>99.0</td>
<td>101.4</td>
<td></td>
</tr>
<tr>
<td>Aggregate Energy Intensity</td>
<td>97.6</td>
<td></td>
<td>68.7</td>
</tr>
</tbody>
</table>
Figure 5. Midwest Decomposition Indexes

<table>
<thead>
<tr>
<th>Index (1998=100)</th>
<th>Real Efficiency</th>
<th>Structural</th>
<th>Aggregate Energy Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>92.7</td>
<td>96.2</td>
<td>89.2</td>
</tr>
<tr>
<td>2002</td>
<td>95.8</td>
<td>84.3</td>
<td>81.7</td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- Blue: 1998
- Red: 2002
- Green: 2006
Figure 6. Northeast Decomposition Indexes
5.0 Regression Analysis

In addition to a decomposition analysis, a regression analysis is undertaken, which examines whether factors such as changes in energy prices are associated (“associated” means a statistically significant relationship) with changes in energy intensity over the 1998 to 2006 time period (with data for the years 1998, 2002, and 2006). This adds to the literature by employing regression analysis, as only a few of the above studies implement regression analysis. A contribution of the regression analysis employed in this study is that it uses more recent data than previous studies. Another contribution to the literature is that this analysis examines between-industry variation, whereas regression studies in the literature tend to focus on within variation.

Unlike other studies examined in the literature review that implement regression analysis, this analysis uses 3-digit NAICS level data for manufacturing industries at the regional and national levels. Bernstein, et. al (2003) used state-level data from 1977-99, which is not broken out by industry; Metcalf (2008) also used state-level data, but his data is broken out into residential, commercial, industrial, and transportation. Wing, et. al (2004) used national-level data for 35 industries over the 1958-96 period, while Wing (2008) used national-level data for 35 industries over the 1958-2000 period.

The regression analysis is conducted at both the national level and regional level.

5.1 National-level Regression Model Methodology

In the national model, the fixed-effects model assumes common slope coefficients and intercepts that vary over industries. The industry specific effects are considered to be fixed.

\[ EI_{it} = \alpha_i + \beta_1 x_{1it} + \beta_2 x_{2it} + \ldots + \varepsilon_{it} \] (1)

EI: Energy intensity

i: Industry

t: Time period

x: time-varying covariates; these include the natural logarithms of the following variables: Price, Avgfloorsp; Avgwage; Empl. The level of Particip and a time trend are also included.

\( \alpha_i \): Industry-specific intercepts
Another common panel data model is the random-effects model. Contrary to the fixed-effects model, the random-effects model assumes the $\alpha_i$ are random.

Its equation is:

$$EI_{it} = \mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + (u_i + \varepsilon_i)$$

$$\alpha_i = \mu + u_i$$

$u_i + \varepsilon_i$: Composite error term

A potential major advantage of the fixed-effects estimator is its consistency when the industry-specific effects are correlated with the time-varying covariates (assuming the time-varying covariates are uncorrelated with the $\varepsilon_i$); the random-effects estimator is inconsistent if the individual specific effects are correlated with the time-varying covariates. A Hausman test can be used to test which model is theoretically preferred. If the model is properly specified and the composite error is uncorrelated with the time-varying covariates (which obviously exclude the NAICS dummy variables), the coefficients estimated by the fixed-effects model should not statistically differ from those estimated by the random-effects model. The Hausman test examines the null hypothesis that the estimates from the two models are not significantly different; the rejection of the null implies the fixed-effects model is preferred over the random-effects model. However, the test is not perfect; for example, some econometric literature indicates fixed-effects coefficient estimates may be especially prone to attenuation bias (bias in the coefficient estimate toward zero) in the case of measurement error (McKinnnish, 2000).

A disadvantage of the fixed-effects model is that it only measures variation within industries, so it does not capture between-industry variation; as the data in Table 3 show, the majority of the variation in the variables occurred between industries, so the fixed-effects model may result in less precise estimates. Conversely, the random-effects model uses variation between industries and variation within industries. When the covariates do not display significant variation over time, implementation of the random-effects model may be required to discover anything about the covariates’ relationship with the covariates (Woodridge 2002). Another disadvantage of the fixed-effects model, is that discarding between-industry variation tends to result in larger standard errors than the random-effects model.
Another alternative is the between model, which simply uses the average of the data for each industry, thus only measuring variation between industries. The reason for including this model in the analysis is to look solely at between-industry variation, since between-industry variation is greater than within-industry variation (Table 3). While the fixed-effects model may exacerbate measurement error, the between model has the advantage in that it reduces the associated bias by averaging out the measurement error (Kennedy 2008).

The between model is as follows:\(^3\):

\[
\bar{EI}_i = \alpha + \beta_1 \bar{x}_{1i} + \beta_2 \bar{x}_{2i} + \ldots + (\alpha_i - \alpha + \bar{\varepsilon}_i)
\]

\(\alpha_i - \alpha + \bar{\varepsilon}_i\): Composite error term

The between model provides consistent estimates in the case that the independent variables are independent of the composite error term. If random-effects model is consistent, the between model is consistent, but less efficient; if the random-effects model is inconsistent, the between model is also inconsistent.

---

\(^3\) The horizontal bar over a variable stands for the mean.
5.2 Regional-level Regression Model Methodology

The regional models are similar to those in Section 5.1, except that they are run at the regional level and variables that are only available at the national level are not included (Partic and Avgflsp).

5.3 National Level Regression Results

The National level regression results are displayed in Table 18. The random-effects and between models find that industries with higher energy prices tend to have lower energy intensities; however, the fixed-effects model does not find a significant price effect, implying that price changes within industries did not significantly affect energy intensity in these industries over the 1998 to 2006 period. This is consistent with the low correlation between the first differences of EI and Price (Table 5), as well as their low degree of within-industry variation relative to between-industry variation (Table 3). An analysis in which data were spaced one year apart, rather than four, might also lead to different results; this would also present itself to testing whether a one-year lag of price is significant.

The fixed-effects model finds most covariates to be insignificant at both the national and regional level, with the exception that changes in average wage are significantly (negatively) associated with changes in energy intensity. One plausible explanation for the lack of significance in the other covariates in the fixed-effects models is that all variables displayed a small amount of variation within industries relative to the amount of variation between industries; since the fixed-effects model omits between-industry variation, it tends to result in higher standard errors of the coefficient estimates. As noted in Section 5.1, fixed-effects models may demonstrate larger standard errors due to ignoring between-industry variation; this effect is exacerbated when the within variation is small. The fixed-effects model has a high R-squared when accounting for industry-specific effects (the industry dummy variables) (0.978), but a low R-squared (0.377) when only accounting for the effects of the within-industry variation.

The random-effects and between models have the following findings in common:

- Industries with higher participation rates in energy efficiency programs tend to have higher energy intensity levels.
- Industries with higher levels of employment tend to have lower energy intensities
- Industries with higher average wages tend to have higher intensities.

Additionally, the between model find that industries with higher average floorspace tend to have lower energy intensities. The random-effects model finds a significant negative time trend for energy intensity, even after accounting for the other covariates.
Table 18. National Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Ln_Price} )</td>
<td>-0.0037 (0.0042)</td>
<td>-0.0106** (0.0046)</td>
<td>-0.0211** (0.0079)</td>
</tr>
<tr>
<td>( \text{Particip} )</td>
<td>0.0060 (0.0217)</td>
<td>0.0390** (0.0195)</td>
<td>0.0946*** (0.0287)</td>
</tr>
<tr>
<td>( \text{Ln_Empl} )</td>
<td>-0.0111 (0.0068)</td>
<td>-0.0144*** (0.0036)</td>
<td>-0.0102** (0.0040)</td>
</tr>
<tr>
<td>( \text{Ln_Avgflsp} )</td>
<td>-0.0026 (0.0040)</td>
<td>-0.0028 (0.0038)</td>
<td>-0.0147** (0.0054)</td>
</tr>
<tr>
<td>( \text{Ln_Avgwage} )</td>
<td>-0.0795*** (0.0253)</td>
<td>0.0187** (0.0084)</td>
<td>0.0186** (0.0074)</td>
</tr>
<tr>
<td>( \text{Year} )</td>
<td>0.0003 (0.0005)</td>
<td>-0.0011*** (0.0005)</td>
<td>-----</td>
</tr>
<tr>
<td>( \text{Constant} )</td>
<td>-0.1255 (0.8860)</td>
<td>2.4076*** (0.7789)</td>
<td>0.2561*** (0.0560)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.978(D(^6))</td>
<td>N/A(^8)</td>
<td>0.914</td>
</tr>
</tbody>
</table>

***99% significance **95% significance *90% significance

---

4 Overall employment is included in place of average employment due to multicollinearity—average employment is very highly correlated with both Avgflsp (correlation coefficient = 0.72) and Avgwage (correlation coefficient = 0.48).
5 A Hausman test rejects the null hypothesis at the 95 percent level, but not at the 99 percent level (p-value = 0.019); therefore, it depends on the desired significance level as to whether the fixed-effects model is preferred to the random-effects model.
6 The R-squared from using each coefficient for each covariate plus each dummy variables for the NAICS.
7 R-squared obtained by only fitting a mean deviated model where the effects of the groups (the dummy variables for NAICS) are assumed to be fixed, so all of the effects for the groups are simply subtracted out from the model and an attempt is not made to quantify their overall effect on the fit of the model. This tends to be similar to the R-squared obtained from running a first difference model, which also only measures within variation.
8 A traditional R-squared measure is not calculated, although Stata reports numerous other R-squared measures, which are not directly comparable.
5.4 Regional Level Regression Results

The regional regression results are displayed in Tables 19 through 22. The fixed-effects models all find a negative coefficient for Ln_Avgwage, and two of these are found to be significant (West and South regions). Other than Ln_Avgwage, no variable is found to be significant by at least two regions. This may be ascribable to factors such as large standard errors resulting from the combination of small within-industry variation and omitting between-industry variation, measurement error, omitted variable bias (not all variables in the national model were available at the regional model), missing data for certain regions/industries, and the fact that the analysis was only for the 1998 to 2006 period (with data only available for three of those years), whereas regression analyses in the literature tend to cover a longer time period and show significant price effects on intensity. The between model for all four regions finds a significant negative coefficient on Ln_price; industries with higher energy prices tend to have lower energy intensities. Only the South region random-effects model finds a significant negative coefficient for Ln_price. Two other regions find a negative Ln_price coefficient, but these are not statistically significant.

Other than the West region, the between and random-effects model each find a significant positive coefficient for Ln_Avgwage. All of the random-effects regional models estimate a significant negative coefficient for Ln_Empl, while two of the regional between models draw the same conclusion. All of the random-effects regional models display a significant negative time trend in energy intensity.
Table 19. West Region Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fixed Effects 9</th>
<th>Random Effects</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln_Price</td>
<td>-0.0006 (0.0013)</td>
<td>-0.0003 (0.0014)</td>
<td>-0.0150** (0.0060)</td>
</tr>
<tr>
<td>Ln_Avgwage</td>
<td>-0.0279* (0.0142)</td>
<td>0.0069 (0.0076)</td>
<td>0.0105 (0.0080)</td>
</tr>
<tr>
<td>Ln_Empl</td>
<td>0.0006 (0.0076)</td>
<td>-0.0090*** (0.0033)</td>
<td>-0.0061 (0.0035)</td>
</tr>
<tr>
<td>Year</td>
<td>0.00003 (0.0003)</td>
<td>-0.0007*** (0.0002)</td>
<td>------</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0599 (0.6819)</td>
<td>1.4641*** (0.4218)</td>
<td>0.0791 (0.0471)</td>
</tr>
</tbody>
</table>

R-squared: 0.978(D10) 0.353(N11) N/A12 0.627

***99% significance  **95% significance  *90% significance

---

9 A Hausman test does not reject the null hypothesis, even at the 90 percent level (p-value = 0.1123); this implies the random-effects model is preferred to the fixed-effects model.

10 The R-squared from using each coefficient for each covariate plus each dummy variables for the NAICS.

11 R-squared obtained by only fitting a mean deviated model where the effects of the groups (the dummy variables for NAICS) are assumed to be fixed, so all of the effects for the groups are simply subtracted out from the model and an attempt is not made to quantify their overall effect on the fit of the model. This tends to be similar to the R-squared obtained from running a first difference model, which also only measures within variation.

12 A traditional R-squared measure is not calculated, although Stata reports numerous other R-squared measures, which are not directly comparable.
Table 20. South Region Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fixed Effects$^{13}$</th>
<th>Random Effects</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln_Price</td>
<td>-0.0055 (0.0059)</td>
<td>-0.0135** (0.0058)</td>
<td>-0.0325** (0.0131)</td>
</tr>
<tr>
<td>Ln_Avgwage</td>
<td>-0.1342** (0.0537)</td>
<td>0.0375** (0.0166)</td>
<td>0.0474** (0.0160)</td>
</tr>
<tr>
<td>Ln_Empl</td>
<td>-0.0221 (0.0170)</td>
<td>-0.0238*** (0.0080)</td>
<td>-0.0151 (0.0091)</td>
</tr>
<tr>
<td>Year</td>
<td>0.0007 (0.0009)</td>
<td>-0.0011* (0.0006)</td>
<td>------</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.6330 (1.7355)</td>
<td>2.3423* (1.2193)</td>
<td>0.1012 (0.1306)</td>
</tr>
</tbody>
</table>

| R-squared  | 0.957(D$^{14}$)      | N/A$^{16}$        | 0.757            |
|           | 0.340(N$^{15}$)      |                 |                  |

***99% significance  **95% significance  *90% significance

---

$^{13}$ A Hausman test does not reject the null hypothesis (p-value = 0.9766); this implies the random-effects model is preferred to the fixed-effects model.

$^{14}$ The R-squared from using each coefficient for each covariate plus each dummy variables for the NAICS.

$^{15}$ R-squared obtained by only fitting a mean deviated model where the effects of the groups (the dummy variables for NAICS) are assumed to be fixed, so all of the effects for the groups are simply subtracted out from the model and an attempt is not made to quantify their overall effect on the fit of the model. This tends to be similar (especially when there are few time periods) to the R-squared obtained from running a first difference model, which also only measures within variation.

$^{16}$ A traditional R-squared measure is not calculated, although Stata reports numerous other R-squared measures, which are not directly comparable.
### Table 21. Midwest Region Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fixed Effects(^{17})</th>
<th>Random Effects</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Ln_Price)</td>
<td>0.0082*</td>
<td>0.0020</td>
<td>-0.0219**</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0046)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>(Ln_Avgwage)</td>
<td>-0.0360</td>
<td>0.0263**</td>
<td>0.0352**</td>
</tr>
<tr>
<td></td>
<td>(0.0236)</td>
<td>(0.0129)</td>
<td>(0.0124)</td>
</tr>
<tr>
<td>(Ln_Empl)</td>
<td>0.0003</td>
<td>-0.0113***</td>
<td>-0.0111***</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0037)</td>
<td>(0.0033)</td>
</tr>
<tr>
<td>Year</td>
<td>0.00003</td>
<td>-0.0006**</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0003)</td>
<td>-----</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0707</td>
<td>1.2546**</td>
<td>0.0613</td>
</tr>
<tr>
<td></td>
<td>(0.8500)</td>
<td>(0.5948)</td>
<td>(0.0597)</td>
</tr>
<tr>
<td>(R)-squared</td>
<td>0.981(D(^{18}))</td>
<td>N/A(^{20})</td>
<td>0.750</td>
</tr>
<tr>
<td></td>
<td>0.210(N(^{19}))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**99% significance  **95% significance  *90% significance

\(^{17}\) A Hausman test rejects the null hypothesis at the 99 percent level (p-value = 0.0068); this implies the fixed-effects model is preferred to the random-effects model.

\(^{18}\) The \(R\)-squared from using each coefficient for each covariate plus each dummy variables for the NAICS.

\(^{19}\) \(R\)-squared obtained by only fitting a mean deviated model where the effects of the groups (the dummy variables for NAICS) are assumed to be fixed, so all of the effects for the groups are simply subtracted out from the model and an attempt is not made to quantify their overall effect on the fit of the model. This tends to be similar (especially when there are few time periods) to the \(R\)-squared obtained from running a first difference model, which also only measures within variation.

\(^{20}\) A traditional \(R\)-squared measure is not calculated, although Stata reports numerous other \(R\)-squared measures, which are not directly comparable.
### Table 22. Northeast Region Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fixed Effects$^{21}$</th>
<th>Random Effects</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln_Price</td>
<td>0.0054 (0.0051)</td>
<td>-0.0038 (0.0043)</td>
<td>-0.0142** (0.056)</td>
</tr>
<tr>
<td>Ln_Avgwage</td>
<td>-0.0306 (0.0358)</td>
<td>0.0263*** (0.0086)</td>
<td>0.0239*** (0.0079)</td>
</tr>
<tr>
<td>Ln_Empl</td>
<td>0.0222 (0.0135)</td>
<td>-0.0179*** (0.0032)</td>
<td>-0.0164*** (0.0031)</td>
</tr>
<tr>
<td>Year</td>
<td>0.0007 (0.0007)</td>
<td>-0.0007* (0.0004)</td>
<td>-----</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.5306 (1.4365)</td>
<td>1.5118** (0.7729)</td>
<td>0.1456*** (0.0406)</td>
</tr>
</tbody>
</table>

R-squared

|                | 0.948(D$^{22}$) | 0.129(N$^{23}$) | N/A$^{24}$ | 0.855 |

***99% significance   **95% significance   *90% significance

---

$^{21}$ A Hausman test rejects the null hypothesis at the 99 percent level (p-value = 0.0002); therefore, this implies the fixed-effects model is preferred to the random-effects model.

$^{22}$ The R-squared from using each coefficient for each covariate plus each dummy variables for the NAICS.

$^{23}$ R-squared obtained by only fitting a mean deviated model where the effects of the groups (the dummy variables for NAICS) are assumed to be fixed, so all of the effects for the groups are simply subtracted out from the model and an attempt is not made to quantify their overall effect on the fit of the model. This tends to be similar to the R-squared obtained from running a first difference model, which also only measures within variation.

$^{24}$ A traditional R-squared measure is not calculated, although Stata reports numerous other R-squared measures, which are not directly comparable.
6.0 Conclusion

This study implemented decomposition and regression analysis to examine energy intensity of the U.S. manufacturing sector over the 1998 to 2006 period, which declined almost 29 percent over the period. The first analysis section decomposed changes in energy intensity into efficiency changes and structural changes, while the second analysis section used regression analysis to examine the association between energy intensity with other variables, such as energy prices and employment.

The conclusions drawn from the decomposition analysis vary, depending on whether a national or regional scheme was used. The national analysis found that 82 percent of the decline in intensity from 1998 to 2006 was due to efficiency and 18 percent was due to structural change. The regional analyses results were mixed, finding efficiency improvements as the main driver of the 45 percent decline in intensity in the West Region and the 31 percent decline in the South Region; conversely, structural change was found to be the main driver of the 18 percent intensity decline in the Midwest Region and the 32 percent intensity decline in the Northeast Region. The structural change component’s dominance in two of the regions points to factors such as international competition from countries with competitive advantages, such as lower wage rates.

The findings of the regression models differ, depending on the type of model used and whether analysis was conducted at the national or regional level. One general conclusion drawn from the regression analysis is that findings may change widely, based on factors such as model choice, aggregation level of the data, and variables and time periods used.

In the national analyses, the random-effects and between models estimate that industries with higher energy prices tend to have lower energy intensities; however, the fixed-effects model did not find a significant price effect. This is consistent with the weak correlations between the first differences of energy price and intensity, contrary to the relatively large cross-sectional correlations between the levels of the two variables. In general, the fixed-effects models find most covariates to be insignificant at both the national and regional level, with the only consistent finding being that changes in average wage are significantly (negatively) associated with changes in energy intensity. One plausible explanation for the lack of significance in the other covariates in the fixed-effects models is that all variables displayed small amount of variation within industries relative to the amount of variation between industries; since the fixed-effects model omits between-industry variation, its coefficient estimates tend to have higher standard errors.

The national random-effects and between models found the following (the term “associated” means a statistically significant relationship):

- Higher participation rates in efficiency programs are associated with higher energy intensities. This may be due to a selection bias, in which industries where higher energy intensities are drawn to such programs more so than industries where energy intensity is relatively low.
- Higher employment levels are negatively associated with intensity.
- Larger average floorspace is associated with lower energy intensities (significant only in between model).
Higher wages are associated with higher energy intensities, contrary to the fixed-effects model.

The random effects model estimates a significant negative time trend.

Higher employment levels are associated with lower energy intensities.

A significant negative price effect is estimated in the regional models in all of the between models, but only one of the random-effects models. This lack of statistical significance of the price variable in the random effects could potentially be due to factors such as omitted variable bias (not all variables in the national model were available at the regional model), missing data for certain regions/industries, and the fact that the analysis was only for the 1998, 2002, and 2006 periods, whereas regression analyses in the literature tended to cover a larger time period and show significant price effects on intensity.

In three of the four regions, the between and random-effects models each find that higher average wages are associated with higher intensities. All of the random-effects regional models and two of the between-regional models find higher employment levels are associated with lower energy intensities. Additionally, each of the random-effects regional models find a statistically significant negative time trend for energy intensity.

An area of potential near-term future research is a similar analysis that appends 2010 data to the 1998, 2002, and 2006 data, once all of the 2010 MECS data are released. This would provide a longer period of analysis from which to draw conclusions for both the decomposition and regression analyses. Additionally, a parallel analysis examining electricity intensity is another possible future area of analysis. A related area of potential future research dealing with energy efficiency is the rebound effect—the idea that improvements in efficiency may lead to increases in demand, which negate some of the energy savings (e.g., more fuel-efficient vehicles may encourage individuals to travel more miles).
References


