Forthcoming article in
Journal of Environmental Management

The Effect of ISO 14001 on Environmental Performance: Resolving Equivocal Findings

Toshi H. Arimura
Faculty of Political Science and Economics
Waseda University
1-6-1 Nishiwaseda, Shinjuku-ku
Tokyo 169-8050, Japan
toshi.arimura@gmail.com

Nicole Darnall*
School of Public Affairs & School of Sustainability
Arizona State University
411 N Central Avenue, Suite 450
Phoenix AZ 85004-0687, USA
ndarnall@asu.edu

Rama Ganguli
Australian Energy Regulator
Level 19, 175 Pitt Street
Sydney, NSW 2000 Australia
ramanuj.g@gmail.com

Hajime Katayama
Faculty of Commerce
Waseda University
1-6-1 Nishiwaseda, Shinjuku-ku
Tokyo 169-8050, Japan
hajikat@gmail.com

* Corresponding Author

The Effect of ISO 14001 on Environmental Performance: Resolving Equivocal Findings

ABSTRACT

Previous research evaluating ISO 14001 certification has shown that it both increases environmental performance and has no effect. We hypothesize that the equivocal findings are due to two factors: institutional pressures arising from differences in regulatory settings across and within countries, and typical methodological approaches for addressing endogeneity. We examine these factors using facility-level data from the United States and Japan for two environmental impacts. After applying Altonji, Elder, and Taber's (2005) method for dealing with the case where there are no potential instruments, we find evidence that the effectiveness of these environmental governance tools varies across countries and type of environmental impact being assessed.

HIGHLIGHTS

- ISO 14001's effectiveness at reducing pollution varies across countries.
- ISO 14001's effectiveness at reducing pollution varies across environmental impact.
- Standard instrumentation approaches to address endogeneity may not be sufficient.
- ISO 14001 may be more effective when there are clear cost-saving opportunities.
- ISO 14001 may be more effective at reducing indirectly regulated impacts.

KEYWORDS: ISO 14001, Environmental management system, Voluntary Environmental Program, Certification program, Environmental performance, Discrete choice model;,Endogeneity

JEL CLASSIFICATION: C35, C36, M14, Q20, Q53, Q58

The Effect of ISO 14001 on Environmental Performance: Resolving Equivocal Findings

1. INTRODUCTION

While traditional government regulations have led to significant reductions in industrial pollution, businesses typically argue that the prescriptive nature of these regulations is costly. Additionally, traditional regulations have been criticized for establishing a threshold of acceptable pollution because facilities typically strive to meet these thresholds but fail to reduce their emissions any further. These limitations and others have caused many governments to promote an alternative environmental policy approach that encourages facilities' voluntary actions to improve the environment beyond that which is required by law. The voluntary approach supplements existing regulations by encouraging cooperation between the regulated community and government officials (Khanna 2001). It is regarded as more flexible and often times less costly than traditional regulations, as facilities have more freedom in choosing their pollution-mitigation strategies. In addition, the voluntary approach does not involve enforcement by regulators, which can reduce regulators' administrative costs. Further, voluntary approaches often encourage facilities to take a broader view of their overall pollution emissions, by emphasizing reductions in unregulated environmental impacts that could lead to greater overall pollution reductions.

One of the most widely used voluntary approaches involves the adoption of the certified environmental management system (EMS) called ISO 14001. By December 2014, 324,148 facilities worldwide had received ISO 14001 certification (International Organisation for Standardisation [ISO] 2015). ISO 14001 is a process standard that grants facilities flexibility in the types of environmental goals they wish to establish. It encourages facilities to systematically manage their environmental impacts by requiring them to implement a series of internal management procedures. Because of its potential to improve the natural environment, some governments have introduced assistance programs that promote facilities' adoption of certified EMSs more broadly (Arimura, Hibiki, and Katayama 2008; USEPA 2010).

Previous assessments of ISO 14001 as a voluntary policy approach have shown mixed results. On the one hand, compared to non-adopters, ISO 14001 adopters operating in the United States (U.S.) (e.g., Potoski and Prakash 2005), Japan (Arimura, Hibiki, and Katayama 2008), and Mexico (Dasgupta, Hettige, and Wheeler 2000) have shown improved environmental performance. More internationalized studies have also shown a positive relationship between ISO 14001 certification and improved environmental performance for manufacturing facilities in Canada, France, Germany, Hungary, Japan, Norway, and the U.S. (Johnstone, Serraville, Scapecchi and Labonne 2007). However, other studies of facilities operating in Mexico (Blackman 2012), the United Kingdom (Dahlstrom, Howes, Leinster and Skea 2003), and the U.S. (King, Lenox, and Terlaak 2005) demonstrate no change in environmental performance or pollution abatement behavior (Ziegler and Rennings 2004), while still other studies in Canada (Barla 2007) and the U.S. (Darnall and Sides 2008) offer inconclusive results.

We argue that there are at least two reasons for the equivocal findings, the first being institutional pressures arising from differences in regulatory settings both across and within countries. Related to cross-country variations, facilities operating within countries with stronger regulatory settings tend to incur more environmental costs which creates incentives for facilities to implement ISO 14001 differently. Related to within country variations, countries utilize different types of environmental laws, with some being more prescriptive and others being more flexible. We suggest that flexibility in the regulatory system creates incentives for ISO 14001 adopters to explore more cost-effective approaches to reduce their environmental impacts, and thus achieve stronger environmental outcomes.

A second reason for the mixed findings may be due to how prior studies have addressed the endogeneity of ISO 14001 adoption. Endogeneity arises because of unobserved facility-specific factors, such as unobservable managers' attitudes towards the environment, and may be positively correlated with both environmental performance and ISO 14001 adoption. As a result, when estimating the relationship between ISO 14001 and environmental performance the error term may be positively correlated with ISO 14001 adoption. The effect of ISO 14001 may be overstated, unless this correlation is adequately controlled for. Endogeneity may also arise due to simultaneity in that facilities with lower environmental performance may be more likely to adopt ISO 14001 in order to improve their performance. In such a case, the effect of ISO 14001 may be understated.

Prior research has addressed issues of endogeneity by relying on an instrumental variable for ISO 14001 adoption (e.g., Arimura, Hibiki, and Katayama 2008; Dasgupta, Hettige, and Wheeler 2000; Johnstone et al. 2007; Potoski and Prakash 2005). The choice of instruments is not an easy task, however. Even if the researcher believes that a variable might be relevant as an instrument, it may not necessarily be independent of the unobserved determinants of the performance and therefore lead to inappropriate conclusions that ISO 14001 is (or is not) effective.

Our research contributes to existing scholarship by exploring these institutional and methodological concerns to determine whether voluntary approaches such as ISO 14001-certification are related to improvements in facilities' environmental performance. Using data from the U.S. and Japan, we compare the facility-level effectiveness of ISO 14001 certification on environmental performance across and within the two countries. We consider two types of regulated impacts: the use of natural resources and air pollutant emissions. In undertaking our estimations, we rely on both the standard instrumentation approach to address endogeneity, in addition to Altonji, Elder, and Taber's (2005) method for dealing with the case where researchers have no potential instruments. After comparing the results from the two estimation approaches, we emphasize the more conservative method to minimize the possibility of drawing inappropriate conclusions about ISO 14001's effectiveness.

Our results offer evidence that the effectiveness of ISO 14001 certification varies across institutional settings – both across and within countries, especially as it relates to the type of environmental impact. Our findings have important implications for the usefulness of this environmental governance tool in that ISO 14001 may be more effective in countries with stronger environmental regulations and for environmental impacts that are governed by more flexible regulation that open the way for greater cost-savings.

2. ISO 14001 AND VARIATIONS IN ENVIRONMENTAL PERFORMANCE

An EMS is a formal set of procedures that defines how a facility will manage its impacts to the natural environment. EMSs require that facilities establish an environmental policy, create quantifiable goals to reduce their environmental impacts, and monitor their environmental progress through systematic auditing and management review (Coglianese and Nash 2001). They are based on a continuous-improvement model that expects firms to periodically revisit and update their environmental improvement goals to ensure that negative environmental impacts are minimized.

While many facilities have employed EMSs for years, ISO 14001 is the first attempt to create an international EMS standard. It requires certification by an independent third-party auditor who helps ensure that the EMS conforms to the ISO 14001 standard. In preparation for certification, a facility must document the procedures that form its EMS. Once certified, the ISO 14001 label indicates that a facility has both implemented a management system that conforms to the ISO 14001 standard and has documentation verifying this conformance (Arimura, Hibiki, and Katayama 2008).

At its core, ISO 14001 is a process standard. Rather than emphasizing actual environmental performance *outcomes*, ISO 14001 emphasizes the *processes* that facilities should undertake to manage their environmental impacts (Khanna 2001). This emphasis allows for variations across facilities' environmental goals and subsequent performance, which opens the possibility that ISO 14001 may not reduce a facility's environmental impacts. This potential is further enhanced by the fact that ISO 14001-certified facilities do not incur regulatory sanctions if they fail to improve their environmental performance. However, by virtue of undergoing certification, facilities are likely to commit to achieving legitimate environmental performance goals (Rondinelli and Vastag 2000), because external verification causes individuals to more seriously consider their group obligations (Olson 1965). External verification also offers a greater degree of accountability and reduces opportunities for participants to behave opportunistically, which can motivate facilities to respond more earnestly to their EMS goals and conform to the ISO 14001 standard (Darnall and Sides 2008).

Studies assessing the performance of ISO 14001 adopters have been mixed, with some studies determining improved environmental performance (Potoski and Prakash 2005; Arimura, Hibiki, and Katayama 2008; Dasgupta, Hettige, and Wheeler 2000), while others have failed to demonstrate environmental improvements (Blackman 2012; Dahlstrom et al. 2003; King, Lenox and Terlaak 2005; Ziegler and Rennings 2004). Still others offer inconclusive results (Barla

2007; Darnall and Sides 2008). We suggest that effectiveness of ISO 14001 certification varies across institutional settings. In particular, we hypothesize that the equivocal findings are due to two factors: country differences and differences in type of environmental impacts.

At the inter-organizational level, institutional pressures arise from external sources such as environmental regulation (Hoffman, 2000). Regulation involves coercive legal mandates for organizations to use pollution control technology, attend to pollution thresholds, and report their pollution emissions to reduce their impact to the natural environment. (Darnall, Henriques and Sadorsky, 2008). Facilities that fail to comply with regulatory requirements risk legal sanction, including losing their operating permits and incurring fines and penalties, which constrains the strategic actions of business, and is the primary reason why organizations implement proactive environmental activities (Hoffman, 1997). We suggest that variations in regulatory settings both across and within countries help explain the equivocal findings about the effectiveness of ISO 14001.

With respect to cross-country variation in the regulatory setting, some countries adopt regulations that are more stringent than others. Facilities operating within more stringent regulatory settings have greater incentives to address environmental concerns. For instance, U.S. firms are not taxed extensively for their use of domestic natural resources, making them less expensive than natural resources in other countries. Additionally, the U.S. is not a signatory to the Kyoto Protocol and so U.S. firms are not pressured in the same way as other firms to minimize their carbon emissions. By contrast, Japanese firms have higher taxes on their use of energy, thus making this natural resource quite expensive comparatively. Moreover, in 2008, Japan committed to achieve a 6% carbon reduction from its 1990 level, thus *indirectly* regulating Japanese facilities' use of natural resources. This regulatory setting creates greater incentives for

Japanese facilities (as opposed to U.S. facilities) to use ISO 14001 to reduce their natural-resource usage. Therefore, we offer the following hypothesis:

Hypothesis 1: ISO 14001's effectiveness at reducing pollution varies across countries.

In addition to cross-country variations in the regulatory setting, variations in regulations within countries may help explain the equivocal findings. For instance, even if natural resources are regulated more in some countries than others, within most developed societies natural resource use is typically regulated using more flexible policy approaches. Such approaches include taxes or performance standards. Facilities therefore have more flexibility to explore cost-effective options when reducing these impacts. For instance, in reducing their natural resource usage, facilities may encourage employees to think differently about their daily routines, including how to reuse their production process water, retrofit lighting and install motion-detecting lights, and utilize process waste so that it becomes an input in the production line. Facilities may also renegotiate their supplier contracts to ensure that a greater proportion of product inputs come from recycled (rather than virgin) materials. By contrast, for pollutants that are governed by more prescriptive regulatory frameworks (e.g., air pollution), facilities are often required to install specific pollution control technologies or a specific change in their operating practices. Such regulation encumbers the sorts of activities a facility may be able to undertake to reduce its pollution, which may make ISO 14001 less effective at reducing these sorts of impacts, at least in the short-run. We therefore offer the following hypothesis:

Hypothesis 2: ISO 14001's effectiveness at reducing pollution varies across environmental impacts.

We suggest that another possible reason why ISO 14001's effectiveness remains unclear is because of methodological concerns related to typical estimation approaches. Prior studies (e.g., Arimura, Hibiki, and Katayama 2008; Dasgupta, Hettige, and Wheeler 2000; Johnstone et al. 2007; Potoski and Prakash 2005) have addressed the endogenous relationship between facilities' ISO 14001 adoption and environmental performance by relying on an instrumental variable. One of the most common instruments (e.g., Arimura, Hibiki, and Katayama 2008; Johnstone et al. 2007) accounts for the availability of a government assistance program for EMS adoption. However, in developing these assistance programs, governments may purposefully target facilities with poor environmental records to encourage them to adopt an EMS. If these facilities subsequently adopt ISO 14001, the availability of an assistance program may not be independent of the unobserved determinants of environmental performance, and therefore lead to inappropriate conclusions about ISO 14001's effectiveness. Other instruments may have similar concerns. For example, Dasgupta, Hettige, and Wheeler (2000) employed a dummy variable for whether the facility was subject to a regulatory inspection. This variable, however, may be correlated with unobserved determinants of compliance if the regulatory agency changes the frequency of inspections for facilities based on their compliance status (Firestone 2002). Similarly, Potoski and Prakash (2005) used emissions as an instrument. Their argument was that a facility's emissions do not directly affect its compliance status, as regulations govern emission-control technologies rather than the level of pollutants. However, unobserved factors such as corporate culture and managers' attitudes towards the environment may affect not only compliance status but also emissions, thus compromising emissions as an instrument. To address this concern, this study therefore uses an alternative method proposed by Altonji, Elder, and Taber (2005) for dealing with the case where researchers have no potential instrument, in addition to standard instrumental variable methods.

3. DATA DESCRIPTION

We draw on data that were collected from a survey developed by the Organisation for Economic Co-Operation and Development (OECD) Environment Directorate and which involved approximately 14 academic researchers and advisory group members located in Canada, France, Germany, Hungary, Japan, Norway, and the U.S. (Johnstone 2007). The survey was designed to gather information on manufacturing facilities' environmental activities within these countries (Johnstone 2007). The survey was vetted by representatives of the OECD's Business and Industry Advisory Committee and pre-tested in Germany, Canada, and Japan, before it was translated into each country's official language and validated for accuracy (Johnstone 2007).

Within Japan, a random sample of 4,757 facility managers was selected from all manufacturing facilities operating within Japan having at least 50 employees. Within the U.S., all 3,746 U.S. manufacturing facilities that reported data to EPA's Toxic Release Inventory were part of the sampling frame, since nearly all manufacturing organizations with at least 50 employees submit data to the inventory. Surveys with an accompanying letter (both OECD and university letterheads) were addressed to each facility's environmental manager and mailed via post in 2003 (Johnstone 2007). Researchers sent two follow-up mailings at three-week intervals to prompt additional responses (Johnstone 2007). The response rate for Japanese facilities was 32% (1,499 respondents). After removing incomplete responses, the Japanese sample contained 1,035 facilities.

The response rate for U.S. facilities was 13% (489 respondents), which is consistent with previous assessments of U.S. organizations' voluntary environmental practices (e.g., Delmas and Keller 2005; Melnyk, Sroufe, and Calantone 2003) where response rates were 11.2 and 10.4%, respectively. After removing incomplete responses, the U.S. sample was 382. Almost half of the OECD sample consisted of small- or medium-sized enterprises (< 250 employees).[1]

Because we use survey data, it is important for us to address issues related to common method bias (CMB). CMB relates to respondents' affective states and their tendency to respond to surveys in a socially desirable way (Podsakoff, MacKenzie, Lee and Podsakoff 2003). Survey research offers two kinds of remedies for CMB, the first being procedural remedies, which occur in the *ex-ante* phase of the research design (Chang, Witteloostuijn, and Eden 2010; Podsakoff et al. 2003). The second remedy is statistical, and takes place in the *ex-post* phase of the research (Chang, Witteloostuijn, and Eden 2010; Podsakoff et al. 2003). Related to procedural remedies, the OECD incorporated several recommended survey-design procedures (e.g., Podsakoff et al. 2003), one being psychological and methodological separation of measurement. To make it appear that the measurement of the dependent variable was not connected with the measurement of the explanatory variables, the OECD utilized six cover stories within the 12-page survey (containing 42 questions), which explained the focus of each specific part of the survey in an effort to separate variables of interest. Additionally, survey questions related to ISO 14001 (page 3) were separated from those related to environmental performance (page 6).

To further reduce the probability of CMB, the survey varied question response formats to engage respondents to a greater degree (Chang, Witteloostuijn, and Eden 2010; Podsakoff et al. 2003). Response formats included Likert scales, yes/no questions, semantic differential scales, and open-ended questions. The OECD reduced item ambiguity by pre-testing the survey. Ambiguity was further diminished by ensuring that target respondents were environmental experts and responsible for their firm's environmental activities (Johnstone 2007). As a final procedural remedy, the survey protected respondent anonymity in order to reduce respondents'

[1] For a more extensive discussion of the OECD data, see Johnstone (2007).

apprehension about the social acceptability of their responses, thus increasing confidence in the quality of the data (Podsakoff et al. 2003).

Related to ex-post statistical remedies that assess CMB, we examine the OECD data using Harman's single-factor test (Podsakoff and Organ 1986). This procedure involves a factor analysis of all the data, and if a single factor emerges accounting for the majority of the covariation between the dependent and independent variables then CMB is a concern (Podsakoff and Organ 1986). The results of our factor analysis lead to multiple factors, suggesting that CMB is less of a concern.

The OECD did not examine non-response bias by comparing early to late responders, and this is a limitation of our study. However, it did evaluate the general distribution of its survey respondents. The OECD found no statistically significant differences when comparing industry representation and firm size of the survey respondents to the distribution of firms in the broader population (Johnstone 2007).

3.1 Dependent Variable

To measure facilities' environmental performance, we consider two types of environmental impact: (1) the use of natural resources (energy, water, etc.), and (2) air pollutant emissions. These impacts are chosen for three reasons. First, findings from previous studies indicate that ISO 14001 helps reduce both facilities' use of natural resources (e.g., Arimura, Hibiki, and Katayama 2008) as well as their air emissions (e.g., Johnstone et al. 2007). Second, these impacts are regulated very differently and therefore exert different pressures on to facilities. Facilities that reduce their natural-resource use typically do so in the absence of a regulatory mandate, and when regulations exist they tend to rely on flexible policies such as taxes and performance standards. As such, facilities can accrue efficiencies related to reductions in product input costs. By contrast, air emissions are prescriptively regulated, and reductions in air emissions beyond regulatory thresholds are often associated with pollution-control investments that can be quite costly. In some instances, facilities that reduce their air pollutants may improve internal efficiencies. However, the efficiency benefits that might accrue from air pollutant reductions are less compelling than that for natural resource reductions because they offer less flexibility in how regulatory compliance is achieved. Examination of these two types of impact therefore provides some insight into whether differential costs and benefits might relate to the effectiveness of ISO 14001.

Third, by examining two types of environmental impacts and using data that were collected equivalently across two different countries, we are in a rare position to assess whether the effectiveness of a voluntary approach differs across countries and also within countries based on regulatory settings. Doing so offers a unique pathway towards understanding the equivocal findings on ISO 14001's effectiveness.

To measure environmental performance, we rely on an OECD survey question that asked environmental managers whether "your facility experienced a change in the environmental impacts per unit of output of its products or production processes in the last three years." Facilities responded by indicating whether they had experienced a "significant decrease," "decrease," "no change," "increase," or "significant increase" in their use of natural resources and in their emission of air pollutants. For both types of environmental impact, responses were aggregated into a binary variable equal to one if the impact had decreased (i.e., an improvement in environmental performance); zero otherwise. Facilities that report "no change" were included in the "no improvement" category for theoretical reasons. The relationship we are interested in is whether ISO 14001 leads to environmental performance improvements, and so our theoretical construct of interest is "improvement." Facilities that report "no change" in their environmental performance do not meet our “improvement” criteria, and including them with facilities that report significant improvements in their environmental performance would misspecify our relationships of interest.

Because of the scaled nature of our dependent variable, a standard ordinal model might also appear suitable. However, it is well known that in nonlinear models such as probit and ordered probit, misclassification can result in biased estimates of the parameters of interest (Hausman et al., 1998; Dustmann and van Soest, 2004). The chance of misclassification is higher with ordinal data than with aggregated binary data, if facilities with equivalent environmental performance respond to the survey question differently. For instance, consider two facilities that reduced their emissions by exactly three percent. Given their identical environmental performance, these facilities should be classified into the same performance category. Consider if one of these facilities instead reports “decrease” while the other reports “significant decrease.” In this instance, an ordinal variable would suffer from misclassification, while a binary variable would not.

Another justification for aggregating our dependent variable is that if the data generating process is a standard ordinal model, then all the underlying parameters except thresholds could be estimated consistently using aggregated binary data[2]. This is important because distinctions between a “significant degrease” as opposed to “decrease” in both the use of natural resources and emissions of air pollutions were less relevant for our purposes. Finally, on a very practical

[2] To see this, consider an ordered probit model for z = 1,2,3,4,5. A latent response variable z^* depends on $\mathbf{x}$ and u, $z^* = \alpha + \boldsymbol{\beta}\mathbf{x} + u$ where $\mathbf{x}$ is a set of covariates and $u|\mathbf{x} \sim N(0,1)$. z^* is associated with z via threshold parameters in the following manner: $z = j$ if $\mu_{j-1} \leq z^* < \mu_j$ where $\mu_0 = -\infty$ and $\mu_5 = \infty$. Usually, μ_1 is normalized to 0 for identification. However, a different normalization, for example, $\mu_2 = 0$ is also possible. Then, the probability that z equals 1 or 2 is $\Pr(z = 1) + \Pr(z = 2) = \Pr(z^* < \mu_2) = \Pr(z^* < 0) = \Pr(u < -\alpha - \boldsymbol{\beta}\mathbf{x}) = \Phi(-\alpha - \boldsymbol{\beta}\mathbf{x})$ where $\Phi(\cdot)$ is the distribution function of the standard normal random variable. Now, consider a variable y where $y = 1$ if $z = 1$ or 2 and $y = 0$ if z = 3, 4, or 5. It follows that $\Pr(y = 1) = \Phi(-\alpha - \boldsymbol{\beta}\mathbf{x})$ and $\Pr(y = 0) = 1 - \Pr(-\alpha - \boldsymbol{\beta}\mathbf{x})$. This means that we can identify α and $\boldsymbol{\beta}$ by simply estimating the probit model with the aggregated binary data y, although threshold parameters (μ_1, μ_3, μ_4) are not identified.

level, the method proposed by Altonji et al. (2005), which we use to address endogeneity, is designed to model a binary dependent variable.

3.2 Key explanatory variable

To assess whether facilities were certified to ISO 14001, we rely on an OECD survey question that asked, “Has your facility acquired any of the following certifications in environmental management?” Facility managers responded whether or not they had certified to ISO 14001. We code this variable equal to one if the facility is certified to ISO 14001; zero otherwise.

3.3 Control variables

Larger facilities are often suggested to have more access to resources and capabilities (Bianchi and Noci, 1998) that may be leveraged towards achieving greater environmental performance (Darnall, Henriques and Sadorsky, 2010). We control for the number of facility employees (logged) as a proxy for facility size. We also account for facility age to control for technology vintage (Russo 2009).

Publicly traded or multinational organizations are more sensitive to brand image, and may encourage them to adopt EMSs (Darnall and Kim 2012). As a consequence, we control for whether or not the facility’s parent company is publicly traded. To account for firms’ financial heterogeneities (Delmas, Etzion and Nairn-Birch 2013) related to their sales, we construct three binary variables to account for whether the facility’s sales have either “increased,” “decreased,” or “remained the same.” The omitted dummy category is whether the facility’s sales have “remained the same.”

Organizations operating in a competitive market are more likely to adopt proactive environmental practices in order to be recognized as being environment-friendly (Khanna and

Speir 2013). To address this issue we account for the facility's market concentration for its most commercially viable product by constructing three dummy variables. These variables account for whether the facility's "number of competitors is greater than 10" (1,0), if its "number of competitors is between 5–10" (1,0), or its "number of competitors is fewer than five" (1,0). "Less than 5 competitors" was our omitted dummy category. We include four dummy variables to control for the facility's market scope since the more export oriented the organization, the higher the benefits it may accrue from its more visible actions to protect the environment (Darnall, Henriques and Sadorsky, 2010; Martín-Tapia, Aragón-Correa and Rueda Manzanares, 2010). They take the value of one if respondents indicate "the scope of the facility's market is global," "regional (neighboring countries)," "national," and "local." The omitted comparison category is the "local market."

Since facilities' environmental practices are influenced by pressures from buyers and consumers (Khanna and Speir 2013), we controlled for where the facility was located within the supply chain, and customers' influence on the adoption of ISO 14001 (Arimura, Hibiki, and Katayama 2008), we use three dummy variables, which equal one (zero, otherwise) if a facility's primary customers are "households" (zero, otherwise), "other facilities within the firm", or "wholesalers/retailers". The omitted dummy category is "other manufacturing firms."

Adopting a quality management system is known to influence facilities' adoption of ISO 14001 (Arimura, Hibiki, and Katayama 2008) because both quality management systems and ISO 14001 are based on complementary continual improvement processes (Ferrón-Vílchez and Darnall 2015). To address this, we include a dummy variable to account for whether or not (1,0) the facility implemented a quality management system.

Technological innovation may also be an important determinant of environmental performance in that facilities with stronger technological innovation may greatly improve their environmental performance. For this reason, technological innovation might be included as a control variable. However, we refrain from doing so in our empirical model, because technological innovation is likely to be correlated with the error term in a way that is similar to ISO 14001 adoption and the inclusion of another endogenous regressor in the model is technically difficult, if not impossible. In addition, the issue of whether or not technological innovation is included does not seem to be relevant to this study, because the objective of this study is to examine the effect of ISO 14001 certification, and because the endogeneity problem that may arise from the omission of technological innovation is addressed with our preferred model.

Descriptive statistics of the variables for both U.S. and Japan facilities are presented in Tables 1 and 2, respectively.

4. METHODOLOGY

To estimate the effect of ISO 14001 certification, we start with a simple probit model with the assumptions that there is no endogeneity of ISO 14001. We then relax this assumption, and assume that we have a valid instrument. We incorporate the traditional instrument—the availability of a government assistance program for EMS adoption—as it was used by prior studies using the same data set as ours (i.e., Johnstone et al. 2007; Arimura, Hibiki, and Katayama 2008). Because of concerns about instrumentation quality, we also rely on Altonji, et al.'s (2005) method for dealing with the case where researchers have no potential instruments. Altonji, et al.'s (2005) method (defined precisely in Section 4.2) is based on the assumption that the degree of selection on unobservables is the same as the degree of selection on the observables.

In our context, the observable and unobservable components of the environmental performance may or may not be associated with ISO 14001 adoption. The assumption means that the degree of association is the same for the observable and unobservable components of the performance.

However, this assumption is not testable, because parameters are just identified by the restriction implied by the assumption (this is just like the assumption that an instrumental variable is not correlated with the error term). Therefore, after comparing the results from the two estimation approaches, we emphasize the results for the more conservative one to minimize the possibility of drawing inappropriate conclusions of ISO 14001's effectiveness.

We assume that a facility will improve its environmental performance if its net benefit from environmental performance improvement is larger than or equal to zero. The net benefit depends on various factors including the adoption of ISO 14001. This is modeled as follows:

$$IMPROVE_i = \mathbf{1}[IMPROVE_i^* \geq 0] \qquad [1]$$

$$IMPROVE_i^* = \delta ISO_i + \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + \varepsilon_{1i}, \qquad [2]$$

where $IMPROVE_i$ is a binary variable that equals to one if facility i made an improvement in the environmental impact, $\mathbf{1}[\ ]$ is an indicator function, $IMPROVE_i^*$ is a latent variable that represents the net benefit from improving the environmental impact, ISO_i is a dummy variable that equals one if the facility adopted ISO 14001 (δ is its coefficient), $\mathbf{x}_{1i}$ is a vector of control variables ($\boldsymbol{\beta}_1$ is the vector of the corresponding coefficients), and ε_{1i} is an idiosyncratic error term.

4.1 Standard probit model

We first assume that ISO_i is exogenous in that it is independent of ε_{1i}. Further we assume that ε_{1i} is standard normally distributed; this results in a probit model where the probability that $IMPROVE_i = 1$ is expressed as

$$\Pr(IMPROVE_i = 1 \mid ISO_i, \mathbf{x}_{1i}) = \Phi(\delta ISO_i + \mathbf{x}'_{1i}\boldsymbol{\beta}_1),$$

where Φ is the distribution function of the standard normal random variable. The effect of interest is the average partial effect (APE) of *ISO*,

$$E_{\mathbf{x}_1}[\Pr(IMPROVE = 1 \mid ISO = 1, \mathbf{x}_1) - \Pr(IMPROVE = 1 \mid ISO = 0, \mathbf{x}_1)],$$

which can be consistently estimated by

$$\frac{1}{N}\sum_{i=1}^{N}[\Phi(\hat{\delta} + \mathbf{x}'_1\hat{\boldsymbol{\beta}}_1) - \Phi(\mathbf{x}'_1\hat{\boldsymbol{\beta}}_1)],$$

where $\hat{\delta}$ and $\hat{\boldsymbol{\beta}}_1$ are estimates of δ and $\boldsymbol{\beta}_1$, respectively.

4.2 Endogeneity of ISO 14001

We now relax the assumption that *ISO* is an exogenous variable. In our context, a standard model that allows for potential endogeneity of *ISO* is a bivariate probit model. In addition to the facility' performance (modeled by equations [1] and [2]), the adoption of ISO 14001 is modeled as

$$ISO_i = \mathbf{1}[ISO_i^* \geq 0] \qquad [3]$$

$$ISO_i^* = \mathbf{x}'_{2i}\boldsymbol{\beta}_2 + \varepsilon_{2i} \qquad [4]$$

where ISO_i^* is a latent variable that represents the net benefit from adopting ISO 14001, $\mathbf{x}_{2i}$ is a set of control variables, and ε_{2i} is an idiosyncratic error term. The error terms (ε_{1i}, ε_{2i}) are assumed to have a bivariate normal distribution with [1, 1, ρ] where the first (second) element is the variance of ε_{1i} (ε_{2i}) and the third element is the correlation between ε_{1i} and ε_{2i}.

Some of the variables in $\mathbf{x}_{2i}$ are usually assumed to be independent of ε_{1i} and hence excluded from $\mathbf{x}_{1i}$. Those variables play a role as "instruments" for identification. Although parameters in a bivariate probit model are identified even without an instrument, it is only through a functional form assumption that identification is achieved. Estimation results that are based solely on a functional form assumption are considered to be unreliable, and therefore researchers typically use instruments to help identification. While the availability of a government assistance program for EMS adoption appears to be the most suitable instrument among all variables in the OECD dataset (Johnstone et al. 2007; Arimura, Hibiki, and Katayama 2008), we have some doubts about its quality for the reasons mentioned earlier.

In the absence of absolutely credible instruments, we estimate the bivariate probit model with $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$. Identification is achieved not by instruments but by a different assumption in an effort to supplement the results derived from the more standard instrumental variable approach. The method we use is the one proposed by Altonji et al. (2005), who formalized the idea of using selection on the observables to study the bias due to unobservables and demonstrate how this information can be incorporated into estimation. In what follows, we briefly explain the method within our context (for details, see Altonji et al. 2005). Hereafter, we drop subscript i for notational simplicity.

Let $IMPROVE^*$ be determined by

$$IMPROVE^* = \delta ISO + \mathbf{w}'\boldsymbol{\gamma}, \quad [5]$$

where $\mathbf{w}$ is the vector containing the full set of variables (both observed and unobserved) determining $IMPROVE^*$, $\mathbf{w} = (w_1, \ldots, w_k)'$, and $\boldsymbol{\gamma}$ is the corresponding coefficient vector, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k)'$. Further, define a dummy variable s_j indicating whether w_j is observed. The observable and unobservable components are then given by

$$\mathbf{x}'\boldsymbol{\gamma}_{\mathbf{x}} = \sum_{j=1}^{k} s_j w_j \gamma_j, \; v = \sum_{j=1}^{k} (1 - s_j) w_j \gamma_j, \quad [6]$$

where $\mathbf{x}$ represents the observable portion of $\mathbf{w}$ with coefficients contained in $\boldsymbol{\gamma}_\mathbf{x}$ (a subvector of $\boldsymbol{\gamma}$) and v is an error component accounting for the unobserved variables in $\mathbf{w}$. Using $\mathbf{x}'\boldsymbol{\gamma}_\mathbf{x}$ and v, we can re-express equation [5] as

$$IMPROVE^* = \delta ISO + \mathbf{x}'\boldsymbol{\gamma}_\mathbf{x} + v, \qquad [7]$$

In this equation, it is highly unlikely that the observed control variables $\mathbf{x}$ are uncorrelated with v. Equation [7] can be re-written as

$$IMPROVE^* = \delta ISO + \mathbf{x}'\boldsymbol{\beta}_1 + \varepsilon_1, \qquad [8]$$

where $\boldsymbol{\beta}_1$ is defined to ensure $COV(\varepsilon_1, \mathbf{x}) = 0$ (i.e., $\mathbf{x}$ is exogenous in equation [8]); $\boldsymbol{\beta}_1$ therefore absorbs both the direct effect of $\mathbf{x}$ on $IMPROVE^*$ (i.e., $\boldsymbol{\gamma}_\mathbf{x}$) and the interaction between $\mathbf{x}$ and v.

Now suppose that the linear projection of ISO^* on $\mathbf{x}'\boldsymbol{\beta}_1$ and ε_1 is given by

$$Proj(ISO^* \mid \mathbf{x}'\boldsymbol{\beta}_1, \varepsilon_1) = \phi_0 + \phi_{\mathbf{x}'\boldsymbol{\beta}_1}\mathbf{x}'\boldsymbol{\beta}_1 + \phi_{\varepsilon_1}\varepsilon_1.$$

Here, $\phi_{\mathbf{x}'\boldsymbol{\beta}_1}$ (ϕ_{ε_1}) measures the impact of the observable (unobservable) component of the *IMPROVE* outcome on ISO 14001 adoption. This can be interpreted as "selection on observables" ("selection on unobservables"). Two conditions on ϕ_{ε_1}, representing extreme assumptions about the degree of selection on unobservables, provide the identifying information for the upper and lower bound estimates of the ISO 14001 effect δ. The first extreme condition (called condition 1) is that the part of *IMPROVE* related to the unobservables has no relationship with ISO^*, i.e., $\phi_{\varepsilon_1} = 0$. If this condition is satisfied, then there is no endogeneity of ISO 14001.

The second extreme condition (called condition 2) is that the part of *IMPROVE* related to the observables has the same relationship with ISO^* as that related to the unobservables, i.e., $\phi_{\varepsilon_1} = \phi_{\mathbf{x}'\boldsymbol{\beta}_1}$. Altonji et al. (2005) motivate the situation where condition 2 is likely to be satisfied. Consider the situation where restrictions on data collection are severe. In such a case, it might be reasonable to interpret the elements of **x** to be a random subset of a very large number of underlying variables **w**. This corresponds to assuming that each s_j in equation [6] is an independent and identically distributed binary random variable equal to one with the same probability for all *j*. Altonji et al. (2005) show that as **w** gets large, condition 2 becomes approximately true for all realizations of s_j.

Although our situation is not exactly what Altonji et al, (2005) suggest, it could be considered an approximation. The data used in this study are a subset of a much larger OECD data set that was assembled to serve a number of research purposes, instead of addressing one research question. Furthermore, in general, data collection is limited by both the knowledge of what the important factors are, as well as the cost and feasibility of gaining information on those factors. Because of these issues, it is reasonable to assume that many elements of **w** are likely to be left out in a relatively random fashion.

Conditions 1 and 2 constrain the degree of selection on unobservables between zero and a ceiling value equal to the degree of selection on the observables:

$$0 \le \phi_{\varepsilon_1} \le \phi_{\mathbf{x}'\boldsymbol{\beta}_1}.$$

Each extreme represents a different identifying assumption for the δ coefficient. Condition 1 corresponds to the restriction that $\rho = 0$ in the bivariate probit specification with $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$. Therefore, an upper bound estimate of the δ coefficient is obtained by estimating a univariate probit model for the performance. Condition 2 can be shown to be operationally equivalent to the following upper bound for ρ

$$\rho = \frac{COV(\mathbf{x}'\boldsymbol{\beta}_1, \mathbf{x}'\boldsymbol{\beta}_2)}{VAR(\mathbf{x}'\boldsymbol{\beta}_1)}.$$

Maximizing the log-likelihood over (δ, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$) subject to the constraint on ρ given above provides a lower bound estimate of the δ coefficient.

We also conduct a sensitivity analysis to reveal how much selection on unobservables is required to eliminate the positive relationship between ISO 14001 and environmental performance. This is done by estimating the bivariate probit model with $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$ and constrain ρ to take values from 0.0 to 0.5 in increments of 0.1. The resulting estimates of the δ coefficient indicate how sensitive the ISO 14001 effect is to the correlation between the unobserved determinants of the performance and the adoption of ISO 14001. The positive relationship between ISO 14001 and environmental performance may be considered "spurious" if it disappears even with a very low level of selection on unobservables.

5. ESTIMATION RESULTS

5.1 Results under the Assumption that ISO 14001 is Exogenous

Model 1 in Table 3 provides the average partial effects (APEs) of *ISO* when *ISO* is assumed to be an exogenous variable. For natural resource usage, the APEs for the U.S. and Japan are estimated to be positive and significant at the 1% levels. According to the results, the former is smaller than the latter; specifically, U.S. adopters of ISO 14001 are 25.9% more likely to reduce their natural resource usage than non-adopters, while Japanese adopters of ISO 14001 are 35.1% more likely to do the same.

ISO 14001 adoption also seems to be associated with the reduction of air pollutant emissions; the APEs for the U.S. and Japan are found to be positive and significant at the 5% levels. In contrast to natural resource usage, however, the APE for U.S. facilities is found to be larger such that U.S. adopters of ISO 14001 are 15.6% more likely to reduce their air pollutant emissions than non-adopters, while Japanese adopters of ISO 14001 are 9.6% more likely to do the same. These results suggest that the effect of *ISO* may differ between the two countries and across the types of impact. The results also demonstrate the importance of controlling for observable effects, especially when estimating the relationship between *ISO* and air pollutant emissions. When we run the same model, but do not control for observable effects, the APEs are 19% ($p<.05$) and 13% ($p<.05$) for the U.S. and Japan, respectively, and thus are overstated by about three percent.[3]

5.2 Correcting for the Potential Endogeneity of ISO 14001: Conventional Method

Next we examine the importance of controlling for the potential endogeneity of ISO 14001. Models 2–4 in Table 3 present the results when we use conventional methods that rely on an instrument. Availability of government assistance programs (*PRGEMP*) is used as an instrument, and thus included in $\mathbf{x}_2$, but not in $\mathbf{x}_1$. In equation [4], the instrument is highly significant both for the U.S. and Japan.[4] This is also the case when we use a linear probability model (i.e., $ISO = \mathbf{x}_2'\boldsymbol{\beta}_2 + u$). Therefore, *PRGEMP* satisfies at least one of the two requirements for a valid instrument in that it is correlated with *ISO* after controlling for other observables.

Model 2 presents the results of the bivariate probit model. For natural resource usage, the APE of *ISO* for the U.S. is not significant at the 5% level. In other words, there is no strong evidence that ISO 14001-certified facilities are more likely to have reduced natural resource usage than non-certified ones. To check the robustness of this result, we also use two different estimation methods: the Rivers and Vuong (1988) two-step method (Model 3) and the two-stage least squares (Model 4). The size and significance of the APE varies substantially between the methods: 54.7% ($p<.01$) for the Rivers and Vuong method (Model 3), while 94.3% (but not significant at $p<.05$) for the two-stage least squares (Model 4). In contrast to the U.S., the APE for Japan is significant at the 1% level and relatively similar across the estimation methods: 43.8% in the bivariate probit (Model 2), 51.8% by the Rivers and Vuong method (Model 3), and 56% by the two-stage least squares (Model 4).

---

[3] The results are available upon request.
[4] The results are available upon request.

For air pollutant emissions, similar results are obtained for U.S. facilities in that the size and significance of the APEs depend heavily on which estimation methods were used. By contrast, all estimation methods suggest that Japanese ISO 14001-certified facilities are no more likely to have reduced air pollutant emissions than non-adopters.

We now ask whether *ISO* is actually an endogenous variable. Results for Models 2–4 provide clear evidence. Related to reductions in natural resource usage for U.S. facilities, the correlation between ε_1 and ε_2 in the bivariate probit model (Model 2) is found to be positive (.586) but not significant at the 5% level. A similar insignificant relationship exists for both Rivers and Vuong's estimation method (Model 3) and the two-stage least squares model (Model 4), suggesting that ISO 14001 is exogenous. Additionally, the correlation between ε_1 and ε_2 is statistically insignificant for Japanese facilities in Models 2–4. A similar pattern is seen for the correlation between ε_1 and ε_2 for reductions in air pollution emissions in that it is statistically insignificant across Models 2–4 for both U.S. and Japanese facilities. These findings indicate that exogeneity cannot be rejected at conventional levels of significance, and that the univariate probit model (Model 1) provides a consistent estimate of the ISO 14001 effect.

5.3 Correcting for the Potential Endogeneity of ISO 14001: Using Altonji et al.'s (2005) Method

The results presented in the previous subsection depend on the assumption that our instrument, *PRGEMP*, is independent of ε_1 in equation [2] (the validity of which cannot be tested). However, as discussed earlier, there is reason to suspect that *PRGEMP* is correlated with ε_1. If so, the finding that ISO 14001 is exogenous could be incorrect. More importantly, all of our estimators that rely on the instrument (i.e., those from the bivariate probit, the Rivers and Vuong method, and the two-stage least squares) would be inconsistent. This possibility motivated us to assume that *PRGEMP* is included in both $\mathbf{x}_1$ and $\mathbf{x}_2$ (i.e., $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$), and to use a different identifying assumption.

Results in Table 4 are obtained by including *PRGEMP* in both $\mathbf{x}_1$ and $\mathbf{x}_2$. Columns (1)–(6) provide the results of a sensitivity analysis, in which we assess how estimates of the ISO 14001 effect change as the correlation parameter (ρ) changes from 0.0 to 0.5. For natural resource usage, when ρ = 0.0 (i.e., ISO 14001 is exogenous), we observe a positive relationship between ISO 14001 and environmental performance for each country. As is expected, the estimated APE decreases when ρ increases. For the U.S., when ρ is assumed to be zero, the APE is 0.243 (p<.01). That is, U.S. ISO 14001-certified facilities are 24.3% more likely to have reduced natural resource usage than non-certified facilities. This result is similar to that obtained in our simple probit model (Model 1, Table 3). However, at ρ = 0.2, the APE of U.S. ISO 14001-certified facilities becomes insignificant at the 5% level. Furthermore, when ρ is 0.4 or larger, the APE becomes negative (though insignificant). The results therefore suggest that U.S. facilities' positive relationship between ISO 14001 and environmental performance measured by natural resource usage is not so robust to moderate levels of the correlation between ε_1 and ε_2.

For Japanese facilities, Table 4 illustrates that when ρ is assumed to be zero, the APE is 0.344, suggesting that Japanese ISO 14001-certified facilities are 34.4% (p<.01) more likely to have reduced natural resource usage than non-certified facilities. The APE declines to 0.287 when ρ = 0.1 and remains positive and significant (p<.05) until ρ = 0.4. Even after ρ = 0.4, the APE remains positive (though insignificant). The correlation between the unobservable components of ISO 14001-certification and reduced natural resource usage would have to be in the range of 0.4–0.5 before the certification effect is insignificant. These findings indicate that

Japanese facilities' positive relationship between ISO 14001 and reductions in natural resource usage is more robust than that for the U.S.

By contrast, for reductions in air pollutant emissions, we find similar results among U.S. and Japanese facilities. When ρ is assumed to be zero, the APEs are 0.144 (p<.05) and 0.088 (p<.05) for the U.S. and Japan, respectively, suggesting that U.S. ISO 14001-certified facilities are 14.4% more likely and Japanese facilities are 8.8% more likely to have reduced air pollutant emissions. However, once ρ becomes 0.1 or larger, the APEs become insignificant for both countries. These results suggest that for both countries the positive relationship between ISO 14001 and reductions in air pollutants is robust to only small levels of the correlation between ε_1 and ε_2.

Column (7) in Table 4 provides the results when we apply Altonji et al.'s (2005) method. This estimation approach provides a conservative lower bound estimate of the effect by using the identifying condition that selection on the unobserved factors is equal to the selection on the observable factors. For natural resource usage, the APE for U.S. facilities is 0.034, but not significant, suggesting that we cannot conclude with confidence that U.S. ISO 14001-certified facilities are more likely to have reduced natural resource usage than non-certified facilities. By contrast, the lower bound estimate of the APE for Japanese facilities is 0.239 (p<.01), suggesting that Japanese facilities' positive relationship between ISO 14001 and reductions in natural resource usage is not spurious.

Related to reductions in air pollutant emissions, the lower bound estimates of the APEs are insignificant for both countries. Hence, we cannot conclude that ISO 14001-certified facilities are more likely to have reduced air pollutant emissions than non-certified facilities.

5.4 Summary of the Results

To summarize our results, related to the natural resource usage of U.S. facilities, the estimated APE of *ISO* differs considerably across the estimation methods, ranging from 0 (Altonji et al.'s method) to 0.55 (Rivers and Vuong's method). In other words, evidence is not robust across the identifying assumptions. Given that the lower bound estimate of APE is not statistically different from zero, the data do not offer strong support for U.S. facilities' positive relationship between ISO 14001 and reductions in natural-resource usage. This is not the case with Japanese facilities, however. No matter which estimation method we use, the APE of *ISO* is positive and significant, ranging from 0.24 (Altonji et al.'s method) to 0.56 (two-stage least squares). We therefore conclude with confidence that ISO 14001 is related to Japanese facilities' reductions in natural resource usage.

Related to air pollutant emissions, for U.S. facilities, we observe a similar pattern to natural resource usage in that the estimated APE of *ISO* ranges from 0 (Altonji et al.'s method) to 0.49 (Rivers and Vuong's method). By contrast, for Japanese facilities, the range of the estimated APEs is rather tight: 0 (Altonji et al.'s method) to 0.10 (standard probit). Importantly, for both countries, the lower bound estimates of the APE are zero. A subsequent sensitivity analysis revealed that the positive relationship between ISO 14001 and reductions in air pollutants disappears in the presence of very small levels of correlation ($\rho = 0.1$). Overall, our results therefore do not support the idea that ISO 14001 certification results in reducing air pollutant emissions[5].

[5] As control variables are of second order importance to the current study, we briefly mention their results obtained by Altonji et al.'s method. We found that when implementing a quality management system, having more employees, having increased their sales, and having more competitors, facilities in Japan are more likely to reduce natural resource usage. Facilities in Japan also tend to reduce air pollutant emissions, when assistant programs are available to them and when their primary customers are wholesalers/retailers. For the U.S., the results are not as clear-cut as those for Japan. Detailed results are available upon request.

Overall, these findings offer support for the idea that variations in institutional pressures from regulation appear to be related to inconsistent findings in ISO 14001's effectiveness. We offer support for Hypotheses 1 and 2, which suggest that equivocal findings are due to two factors: country differences and differences in type of environmental impacts.

6. DISCUSSION AND CONCLUSION

This research adds to a growing body of literature that addresses the increasing prominence of voluntary environmental governance approaches and their role as supplements to traditional regulation. The most recognized of these approaches, ISO 14001, has demonstrated mixed performance outcomes, with some prior studies indicating that certification increases environmental performance while other studies show no effect.

Our results provide a basis for understanding the equivocal findings in the scholarly literature, and ISO 14001 as a voluntary policy approach. More specifically, our results are consistent with the possibility that institutional pressures arising from differences in regulatory settings across countries and within countries are related to ISO 14001 outcomes. Related to differences in regulatory settings across countries, Japanese (but not U.S.) ISO 14001-certified facilities are more likely to have reduced natural resource usage than non-certified facilities, most likely because natural resources are more highly regulated in Japan. Because of higher taxes, energy in Japan all tend to be more expensive than the same natural resources in the U.S. Additionally, Japanese facilities' natural resource inputs are indirectly regulated by way of the Kyoto Protocol. This setting creates greater incentives for the conservation of natural resources in Japan that can be achieved since ISO 14001 is a process standard, and facilities have significant flexibility in developing cost-effective pollution mitigation strategies. By contrast, U.S. facilities have access to cheaper natural resources that are not indirectly regulated because the U.S. is not signed onto the Kyoto Protocol. U.S. facilities therefore have fewer pressures to reduce their natural resources.

Additionally, for both countries, ISO 14001-certified facilities are no more likely to reduce their air pollution emissions than non-certified ones. These findings are likely due to within country variations in the regulatory setting. Existing laws in both countries require manufacturing facilities to reduce air pollution by prescriptive regulatory approaches such as installing specific pollution-control technologies or changing their operating practices in a specific way. Such prescriptions often require that additional reductions in air pollution involve renegotiating the facility's air pollution permit, or installing expensive control technology. This situation may discourage facilities (in both countries) from using ISO 14001 to reduce their air pollution emissions. If so, the prescriptive regulatory settings would appear to constrain the applicability of voluntary approaches such as ISO 14001, making it less effective (at least in the short-run). By contrast, regulation of natural resources is typically governed by a more flexible policy approach that is based on taxes or performance goals. Facilities that reduce their natural resource usage typically do so by improving their operational efficiencies and reducing product input costs, which can yield significant cost savings. In spite of these cost saving opportunities, U.S. facilities were not persuaded to reduce their natural resource usage. Rather, we attribute country-level heterogeneities related to Japan's resource scarcity and indirect regulation of natural resources to its facilities' greater emphasis on reducing natural resource usage.

While our results might indicate some constraints on the effectiveness of voluntary environmental approaches, even if an ISO 14001-certified facility does not directly improve its environmental performance, it is important to note that ISO 14001 adoption might affect the environmental actions of other organizations and especially those in the supply chain (Arimura, Darnall, and Katayama 2011). Indeed, ISO 14001-certified facilities are 40% more likely to assess their suppliers' environmental performance and 50% more likely to require that their suppliers undertake specific environmental practices (Arimura, Darnall, and Katayama 2011), which may create compelling incentives for these suppliers to reduce their overall environmental impacts. The total positive externalities that might accrue from this voluntary environmental governance approach therefore may be quite significant even if the adopting facility only achieves modest environmental improvements.

In sum, while many governments promote voluntary approaches to improve the environment beyond that which is required by law, there is still much to learn about this non-regulatory approach. Our results offer important information about the relationship between ISO 14001 certification and environmental performance, and largely support our idea that studies assessing the effectiveness of these environmental governance tools should consider institutional variations across countries and idiosyncrasies related to the type of environmental impact being assessed. Additionally, scholars should go beyond typical approaches for addressing endogeneity. Overall, our findings point to the effectiveness of ISO 14001 as a voluntary governance tool, and suggest that eco-certification might be particularly useful to reduce environmental impacts that are indirectly regulated and costly, as opposed to costly impacts that are highly regulated or impacts that are generally unregulated and inexpensive.

7. ACKNOWLEDGEMENTS

The authors thank the Organisation for Economic Cooperation and Development for funding a portion of this study. The views expressed in this paper are those of the authors, and do not necessarily reflect the views or policies of Australian Energy Regulator. All conclusions and any errors are the responsibility of the authors.

References

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113:151–184.

Arimura, T. H., N. Darnall, and H. Katayama. 2011. "Is ISO 14001 a Gateway to More Advanced Voluntary Action? The Case of Green Supply Chain Management." *Journal of Environmental Economics Management* 61:170–182.

Arimura, T. H., A. Hibiki, and H. Katayama. 2008. "Is a Voluntary Approach an Effective Environmental Policy Instrument? A Case for Environmental Management Systems." *Journal of Environmental Economics Management* 55:281–295.

Barla, P. 2007. "ISO 14001 Certification and Environmental Performance in Quebec's Pulp and Paper Industry." *Journal of Environmental Economics Management* 53:291–306.

Bianchi, R. and Noci, G. 1998. "Greening' SMEs' Competitiveness." *Small Business Economics* 11:269-281.

Blackman, A. 2012. "Does Eco-Certification Boost Regulatory Compliance in Developing Countries? ISO 14001 in Mexico." *Journal of Regulatory Economics* 42:242–263.

Chang, S.J., A. van Witteloostuijn, and L. Eden. 2010. "From the Editors: Common Method Variance in International Business Research." *Journal of International Business Studies* 41:178–184.

Coglianese, C., and J. Nash, eds. 2001. *Regulating from the Inside: Can Environmental Management Systems Achieve Policy Goals?* Washington, DC: Resources for the Future.

Dahlstrom, K., C. Howes, P. Leinster, and J. Skea. 2003. "Environmental Management Systems and Company Performance: Assessing the Case for Extending Risk-based Regulation."

European Environment 13:187–203.

Darnall, N., and D. Edwards Jr. 2006. “Predicting the Cost of Environmental Management System Adoption: The Role of Capabilities, Resources, and Ownership Structure.” *Strategic Management Journal* 27:301–320.

Darnall, N., I. Henriques, and P. Sadorsky. 2008. “Do environmental management systems improve business performance in an international setting?” *Journal of International Management* 14: 364-376.

Darnall, N., I. Henriques, and P. Sadorsky. 2010. “Adopting Proactive Environmental Practices: The influence of Stakeholders and Firm Size.” *Journal of Management Studies* 47:1072-1094.

Darnall, N. and Y. Kim. 2012. “Which types of environmental management systems are related to greater environmental improvements?” *Public Administration Review* 72:351-365.

Darnall, N., and S. Sides. 2008. “Assessing the Performance of Voluntary Environmental Programs: Does Certification Matter?” *Policy Studies Journal* 36:95–117.

Delmas, M., D. Etzion, and N. Nairn-Birch. 2013. “Triangulating Environmental Performance: What do Corporate Social Responsibility Ratings Really Capture?” *Academy of Management Executive* 27:255-267.

Dasgupta, S., H. Hettige, and D. Wheeler. 2000. “What Improves Environmental Compliance? Evidence from Mexican Industry.” *Journal of Environmental Economics and Management* 39:39–66.

Delmas, M., and A. Keller. 2005. “Free Riding in Voluntary Environmental Programs: The Case of the U.S. EPA WasteWise Program.” *Policy Sciences* 38:91–106.

Dustmann, C., and A. van Soest. 2004. “An Analysis of Speaking Fluency of Immigrants Using

Ordered Response Models with Classification Errors." *Journal of Business and Economic Statistics* 22:312-321.

Ferrón-Vílchez, V, and N. Darnall. 2014. "Two are Better than One: The Link between Management Systems and Business Performance." *Business Strategy and the Environment* DOI: 10.1002/bse.1864

Firestone, J. 2002. "Agency Governance and Enforcement: The Influence of Mission on Environmental Decisionmaking." *Journal of Policy Analysis and Management* 21:409–426.

Goocher, B. E. 1965. "Effects of Attitudes and Experience on the Selection of Frequency Adverbs." *Journal of Verbal Learning and Verbal Behavior* 4: 193-195.

Hausman, J. A., J. Abrevaya, F.M. Scott-Morton. 1998. "Misclassification of a Dependent Variable in Discrete Response Setting." *Journal of Econometrics*, 87:239-269.

Hoffman, A. 2000. Competitive Environmental Strategy: A Guide to the Changing Business Landscape. Island Press, Washington, D.C.

Hoffman A. 1997. From Heresy to Dogma: An Institutional History of Corporate Environmentalism. New Lexington Press, San Francisco.

International Organisation for Standardisation. 2015. *The ISO Survey of Management System Standard Certifications 2014.* Geneva, Switzerland: ISO.

Johnstone, N., ed. 2007. *Environmental Policy and Corporate Behaviour*. Cheltenham, UK: OECD, Edward Elgar Publishing.

Johnstone, N., C. Serraville, P. Scapecchi, and J. Labonne. 2007. "'Many a Slip 'twixt the Cup and the Lip': Direct and Indirect Public Policy Incentives to Improve Corporate Environmental Performance." In *Environmental Policy and Corporate Behaviour*, ed. N. Johnstone. Cheltenham, UK: OECD, Edward Elgar Publishing.

Khanna, M. 2001. “Non-mandatory Approaches to Environmental Protection.” *Journal of Economic Surveys* 15:291–324.

Khanna, M and C Speir. 2013. “Motivations for Proactive Environmental Management.” *Sustainability*, 5:2664-2692.

King, A., M. Lenox, and A. Terlaak. 2005. “The Strategic Use of Decentralized Institutions: Exploring Certification with the ISO 14001 Management Standard.” *Academy of Management Journal* 48:1091–1106.

Martín-Tapia, I., J. A. Aragón-Correa, and A. Rueda-Manzanares, A. 2010 “Environmental Strategy and Exports in Medium, Small and Micro-enterprises.” *Journal of World Business* 45:266-275.

Melnyk, S. A., R. P. Sroufe, and R. L. Calantone. 2003. “Assessing the Impact of Environmental Management Systems on Corporate and Environmental Performance.” *Journal of Operations Management* 21:329–351.

Olson, M. 1965. *The Logic of Collective Action.* Cambridge, MA: Harvard University Press.

Podsakoff, P. M., S. B. MacKenzie, J. Y. Lee, and N. P. Podsakoff. 2003. “Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies.” *Journal of Applied Psychology* 88 (5): 879–903.

Podsakoff, P. M., and D. W. Organ. 1986. “Self-Reports in Organizational Research: Problems and Prospects.” *Journal of Management* 12:531–544.

Potoski, M., and A. Prakash. 2005. “Covenants with Weak Swords: ISO 14001 and Facilities’ Environmental Performance.” *Journal of Policy Analysis and Management* 24:745–769.

Rivers, D., and Q. H. Vuong. 1988. “Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models.” *Journal of Econometrics* 39:347–366.

Rondinelli, D. A., and G. Vastag. 2000. “Panacea, Common Sense, or Just a Label? The Value of ISO 14001 Environmental Management Systems.” *European Management Journal* 18:499–511.

Russo, Michael V. 2009. “Explaining the Impact of ISO 14001 on Emission Performance: A Dynamic Capabilities Perspective on Process and Learning.” *Business Strategy and the Environment* 18:307-319.

Schaeffer, N. C. 1991. “Hardly Ever or Constantly? Group Comparisons Using Vague Quantifiers.” *Public Opinion Quarterly* 55: 395-423.

U.S. Environmental Protection Agency. 2010. “Environmental Management System/ISO 14001 – Frequently Asked Questions.” Washington, DC: U.S. EPA, Office of Water. Available at http://water.epa.gov/polwaste/wastewater/Environmental-Management-System-ISO-14001-Frequently-Asked-Questions.cfm (last accessed September 13, 2014).

Ziegler, A., and K. Rennings. 2004. “Determinants of Environmental Innovations in Germany: Do Organizational Measures Matter? A Discrete Choice Analysis at the Firm Level.” *ZEW Discussion Paper* 04-30. Center for European Economic Research.

Table 1: Summary Statistics for the U.S.*

Variable	ISO 14001 adopters		ISO 14001 non-adopter	
	Mean	***Std. Dev.***	***Mean***	***Std. Dev.***
IMPROVE (the use of natural resource)	0.74	0.44	0.48	0.50
IMPROVE (the air pollutant emissions)	0.70	0.46	0.50	0.50
Number of employees	542.34	640.11	400.82	549.52
Facility age	45.59	25.83	47.38	35.62
Facility belongs to publicly traded firm	0.64	0.48	0.39	0.49
Facility sales have increased	0.23	0.43	0.32	0.47
Facility sales have declined	0.44	0.50	0.40	0.49
Primary customer is other manufacturing firms	0.78	0.42	0.49	0.50
Primary customer is wholesaler/retailer	0.15	0.36	0.42	0.49
Primary customer is household consumers	0.03	0.16	0.06	0.24
National market scope	0.19	0.40	0.36	0.48
Regional (neighboring countries) market scope	0.08	0.28	0.08	0.28
Global market scope	0.71	0.46	0.50	0.50
Firm has less than 5 market competitors	0.26	0.44	0.27	0.44
Firm has between 5–10 market competitors	0.38	0.49	0.40	0.49
Quality management system	0.95	0.23	0.60	0.49
Availability of government EMS assistance programs	0.53	0.50	0.35	0.48
Food	0.01	0.12	0.10	0.30
Textiles	0.03	0.16	0.02	0.15
Wood	0.01	0.12	0.08	0.27
Pulp	0.04	0.20	0.05	0.21
Chemicals	0.36	0.48	0.25	0.44
Other non-metallic mineral	0.01	0.12	0.06	0.23
Basic metals	0.27	0.45	0.26	0.44
Machinery	0.12	0.33	0.12	0.33
Transport equipment	0.12	0.33	0.05	0.22
Recycling	0.01	0.12	0.01	0.11

* n = 73 for ISO 14001 adopters; n = 309 for ISO 14001 non-adopters.

Table 2: Summary Statistics for Japan*

Variable	ISO 14001 adopters		ISO 14001 non-adopter	
	Mean	***Std. Dev.***	**Mean**	***Std. Dev.***
IMPROVE (the use of natural resource)	0.72	0.45	0.37	0.48
IMPROVE (the air pollutant emissions)	0.45	0.50	0.33	0.47
Number of employees	558.03	1962.53	141.03	179.51
Facility age	39.57	23.00	42.31	23.66
Facility belongs to publicly traded firm	0.23	0.42	0.04	0.21
Facility sales have increased	0.25	0.43	0.20	0.40
Facility sales have declined	0.50	0.50	0.56	0.50
Primary customer is other manufacturing firms	0.71	0.45	0.62	0.49
Primary customer is wholesaler/retailer	0.15	0.36	0.28	0.45
Primary customer is household consumers	0.07	0.25	0.08	0.27
National market scope	0.60	0.49	0.70	0.46
Regional (neighboring countries) market scope	0.01	0.12	0.01	0.11
Global market scope	0.33	0.47	0.13	0.33
Firm has less than 5 market competitors	0.30	0.46	0.28	0.45
Firm has between 5–10 market competitors	0.38	0.49	0.34	0.47
Quality management system	0.93	0.26	0.68	0.47
Availability of government EMS assistance programs	0.30	0.46	0.13	0.34
Food	0.04	0.19	0.14	0.35
Textiles	0.03	0.16	0.06	0.24
Wood	0.01	0.11	0.02	0.15
Pulp	0.07	0.25	0.07	0.25
Chemicals	0.15	0.36	0.13	0.34
Other non-metallic mineral	0.00	0.07	0.03	0.18
Basic metals	0.15	0.36	0.22	0.41
Machinery	0.41	0.49	0.24	0.43
Transport equipment	0.13	0.33	0.06	0.24
Recycling	0.01	0.12	0.02	0.13

* n = 405 for ISO 14001 adopters; n = 630 for ISO 14001 non-adopters.

Table 3: Average Partial Effect of ISO 14001 – No Instrument and Conventional Instrumentation Methods

Environmental Impacts by Country and Exogeneity Test	No Instrument	Instrumentation – Conventional Estimation Methods		
	Model 1 Probit	*Model 2* Bivariate Probit	*Model 3* Rivers & Vuong (1998) Two-step Method	*Model 4* Two-stage Least Squares
Reduce Natural Resource Usage				
U.S. (N = 382)	0.259**	−0.104	0.547**	0.943
	(0.062)	(0.496)	(0.155)	(0.522)
Exogeneity Test		0.586	−1.812	−0.696
		(0.684)	(1.194)	(0.452)
Japan (N =1035)	0.351**	0.438**	0.518**	0.560**
	(0.033)	(0.090)	(0.129)	(0.200)
Exogeneity Test		−0.178	−0.599	−0.215
		(0.186)	(0.548)	(0.198)
Reduce Air Pollution Emissions				
U.S. (N = 365)	0.156*	−0.385	0.492*	0.725
	(0.069)	(0.293)	(0.234)	(0.554)
Exogeneity Test		0.908	−1.554	−0.583
		(0.545)	(1.333)	(0.511)
Japan (N = 783)	0.096*	0.207	0.306	0.325
	(0.040)	(0.113)	(0.176)	(0.220)
Exogeneity Test		−0.197	−0.625	−0.233
		(0.192)	(0.575)	(0.221)

Note: This table presents the average partial effects of ISO 14001 on the probability that *Reduce Natural Resource Usage* = 1 (i.e., the facility decreased natural resource usage) and on the probability that *Reduce Air Pollution Emissions* = 1 (i.e., the facility decreased air pollutants emission). For Models 1, 2, and 4, robust standard errors are presented in parentheses. For Model 3, bootstrap (500 replications) standard errors are in parentheses. ** and * indicate significance at the 1% and 5% level, respectively. "Exogeneity test" for Model 2 presents the estimate of the correlation between ε_1 and ε_2 (i.e., ρ). Under the null that *ISO* is exogenous, ρ is equal to zero. "Exogeneity test" for Models 3 and 4 present the coefficient on $\hat{v}$ in the linear (probit) regression of *IMPROVE* on $\hat{v}$, *ISO*, and $\mathbf{x}_1$, where $\hat{v}$ is the residual obtained by regressing *ISO* on $\mathbf{x}_2$. Under the null that *ISO* is exogenous, the coefficient on $\hat{v}$ is zero.

Table 4: Average Partial Effect of ISO 14001 –
Different Assumptions on the Correlation of Disturbances

Environmental Impacts by Country	Constraints on the Correlation of the Disturbances, ε_1 and ε_2						Altonji et al.'s (2005) Method	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	ρ = 0.0	ρ = 0.1	ρ = 0.2	ρ = 0.3	ρ = 0.4	ρ = 0.5	$\rho = \frac{COV(\mathbf{x}'\boldsymbol{\beta}_1, \mathbf{x}'\boldsymbol{\beta}_2)}{VAR(\mathbf{x}'\boldsymbol{\beta}_1)}$	
	APE	***APE***	***APE***	***APE***	***APE***	***APE***	***APE***	ρ
Reduce Natural Resource Usage								
U.S. [N = 382]	0.243**	0.187**	0.128	0.065	-0.000	-0.068	0.034	0.403
	(0.063)	(0.066)	(0.069)	(0.069)	(0.075)	(0.066)	(0.233)	(0.348)
Japan [N = 1035]	0.344**	0.287**	0.225**	0.159**	0.090*	0.017	0.239**	0.202**
	(0.034)	(0.035)	(0.010)	(0.037)	(0.036)	(0.035)	(0.059)	(0.053)
Reduce Air Pollution Emissions								
U.S. [N = 365]	0.144*	0.085	0.024	-0.039	-0.102	-0.164*	0.083	0.196*
	(0.070)	(0.072)	(0.073)	(0.080)	(0.068)	(0.064)	(0.469)	(0.805)
Japan [N = 783]	0.088*	0.028	-0.031	-0.089*	-0.153**	-0.203**	0.047	0.133*
	(0.040)	(0.040)	(0.039)	(0.038)	(0.036)	(0.033)	(0.068)	(0.062)

Note: This table presents the average partial effects of ISO 14001 on the probability that *Reduce Natural Resource Usage* = 1 (i.e., the facility decreased natural resource usage) and on the probability that *Reduce Air Pollution Emissions* = 1 (i.e., the facility decreased air pollutants emission) in the bivariate probit models where the correlations in the latent error terms are constrained to 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, and $COV(\mathbf{x}'\boldsymbol{\beta}_1, \mathbf{x}'\boldsymbol{\beta}_2)/VAR(\mathbf{x}'\boldsymbol{\beta}_1)$. Robust standard errors are in parentheses. ** and * indicate significance at the 1% and 5% level, respectively.