**Causal mediation in educational research: An illustration using international
assessment data**

Daniel H. Caro

Oxford University Centre for Educational Assessment

University of Oxford

Abstract

This paper applies the causal mediation framework proposed by Kosuke Imai and colleagues (Imai, Keele, & Tingley, 2010) to educational research by examining the causal mediating role of early literacy activities in parental education influences on reading performance. The paper argues that the study of causal mediation is particularly relevant in educational settings because learning outcomes result from complex interactions involving multiple actors. The analyses use retrospective longitudinal data from the Progress in International Reading Literacy Study (PIRLS) 2011. The value of international assessment data for producing evidence of causation is discussed. The paper aims to contribute to the study of causation with observational data from educational studies.

Causation in educational research

Recent years have seen an increased interest among educationalists in statistical techniques, such as instrumental variables, regression discontinuity, differences-in-differences, and propensity score matching to establish causation with observational data (Kaplan, 2009; Murnane & Willett, 2011; Schlotter, Schwerdt, & Woessmann, 2011). This is reflected in a number of publications and workshops taking place in North America and Europe. For the most part, it has been economists who have taken the lead in this area (Angrist & Pischke, 2009; Schlotter, Schwerdt, & Woessmann, 2011). The proposed techniques resemble randomized experiments and can be very effective at identifying *causal effects* (e.g., instrumental variables and regression discontinuity), provided that certain stringent assumptions are met. However, they represent one view of causation, even within the econometrics field. And critics argue that the absence of explicit theoretical models and focus on treatment effects provides merely a black-box approach to causation that is not helpful for understanding the mechanisms at work for causal effects (Heckman, 2008). For example, a regression discontinuity approach based on exogenous maximum class size policies is able to show that class size has a negative effect on academic achievement (Angrist & Lavy, 1999), but cannot explain why and how it is that class size reductions promote academic achievement.

Arguably, the identification of causal effects is of little use for educationalists if the *causal mechanisms* are not well understood. This is particularly relevant in educational settings because causal effects arise from complex family and institutional processes that involve interactions between a number of actors (e.g., students, parents, teachers, principals, governments) and contexts (e.g., home, pre-school, school, neighbourhood). The design of educational interventions informed by causal evidence requires a careful consideration of these processes. Since atheoretical applications of econometric techniques cannot capture the

complex mechanisms and interactions governing causal effects, their contribution to educational policy and theory is limited (Kaplan, 2009; Murnane & Willett, 2011). In fact, it is debatable that any causation evidence can be established without an understanding of the mechanisms at work.

Other techniques, such as mediation analysis, seem more appropriate for identifying the mechanisms underlying observed relationships (Baron & Kenny, 1986; MacKinnon, 2008). Mediation analysis reveals these mechanisms by evaluating the role of mediating variables (i.e., indirect effects) in treatment effects. However, it is not clear that mediation results can be interpreted in causal terms (Imai, Keele, & Tingley, 2010; Muthén, 2011; VanderWeele & Vansteelandt, 2009). Recently, researchers have been working towards the development of a causal framework for mediation analysis (Muthén, 2011; Imai, Keele, Tingley, & Yamamoto, 2011; Imai & Yamamoto, 2013). And already causal mediation techniques have been implemented in the R software (Tingley, Yamamoto, Keele, & Imai, 2013). This paper illustrates an application of causal mediation analysis with international assessment data.

Causation and international assessments

International assessments of student achievement like the Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS) have played an important role in educational research and policy (e.g., Baird, et al., 2011; Grek, 2009). But researchers have also recognised the design and theoretical limitations of these studies for providing evidence of causation ([citation removed for blind review]; Gustafsson, 2013; Leung & van de Vijver, 2008).

On the one hand, different sources of bias in cross-sectional designs prevent researchers from interpreting associations as *causal effects* (Gustafsson, 2013). For example,

observed negative associations between parental involvement and academic achievement in PIRLS and PISA ([citation removed for blind review]) are probably not an indication that parental involvement produces lower achievement but rather that parents react to lower achievement with greater involvement (i.e., endogeneity or reverse causality bias). Similarly, reported negative associations between academic achievement and time spent on homework likely reflect endogeneity bias, since parents and teachers tend to assign more homework to students with lower achievement (Gustafsson, 2013). Additionally, studies are affected by omitted variable bias, that is, unobserved variables that explain observed associations.

On the other hand, the lack of theoretical foundations in the research programme of international assessments limits the possibilities for studying *causal mechanisms*. For example, the extended use of indicators based on a small number of items, poorly grounded in theory, and not equally valid across cultures, such as family socioeconomic status (SES) or school climate, limits the possibility of determining the causal mechanisms behind relationships (Artelt, 2005; [citation removed for blind review]; Hopfenbeck & Maul, 2011).

Attempts to investigate causal effects and causal mechanisms

Some studies have tried to overcome these limitations by using statistical techniques and study designs that resemble randomized experiments to produce evidence of *causal effects*. For example, using trend data from different international assessments and a differences-in-differences design, Hanushek and Wößmann (2006) studied the effect of school tracking on academic achievement. In another study, Wößmann and West (2006) used data from TIMSS and instrumental variables techniques to evaluate the effect of class size on student performance. Also, using data from PIRLS and regression discontinuity models, Martin, Mullis, and Foy (2011) analysed the effect of years of schooling on reading performance. And, Dronkers and Avram (2009) used PISA data and propensity score

matching to compare private-dependent schools and public schools whilst adjusting for their intake selectivity.

Gustafsson (2013) has proposed an analytical approach consisting of data aggregation and longitudinal analysis to remove endogeneity bias and omitted variable bias (see also Van Damme, Liu, Vanhee, & Pustjens, 2010). He has illustrated this approach with TIMSS data by analyzing effects of class size and time spent on homework on academic achievement. Gustafsson (2013) argues that endogeneity bias (e.g., the influence of student performance on time spent of homework) is solved by aggregating data at the country level, because the mechanisms causing reverse causality at the student level need not be present at the higher levels of observation. And omitted variable bias (e.g., class size is related to unobserved quality of teaching) is removed by analysing changes over time in country-level aggregated variables, that is, by removing countries' fixed effects. Analyses at the country level, however, typically rely on very small samples and might be affected by aggregation bias and ecological fallacy.

Researchers have also worked towards improving the theoretical foundations of international assessments and thereby the possibilities for identifying *causal mechanisms*. For example, [citation removed for blind review] used PISA and PIRLS data to develop cultural, social, and economic capital constructs well-grounded in theory and evaluate their cross-cultural validity. Studies have also evaluated the validity and cross-cultural generalisability of students' motivation (Artelt, 2005) and approaches to learning (Marsh, Hau, Artelt, Baumert, & Peschar, 2006) constructs. These studies provide benchmarks for investigating the mechanisms governing observed associations with the data. But due to the large-scale nature of international assessments, it is in general difficult to comprehensively evaluate psychological and sociological constructs. Similarly, it is costly to include and collect information on variables unrelated to academic achievement in conceptual frameworks,

which sometimes make useful instrumental variables for capturing exogenous variation in treatment effects.

Also, in a recent study with PISA data, [citation removed for blind review] applied concepts from educational effectiveness research to international assessment studies. The effectiveness concept measures performance in relative terms, that is, by taking into account non-malleable factors for educational systems such as SES at the different levels (i.e., student, school, and country) and prior academic achievement. [citation removed for blind review] produced indicators of effectiveness performance at the education system level using PISA trend data (2000-2009). This approach offers a new classification of education systems based on effectiveness rather than on absolute performance and equips educational policymakers with a complementary view of the performance of education systems that lays the foundation for the study of policies, structures, and effectiveness-enhancing factors.

Comparison with randomized and longitudinal designs

Researchers have also argued that international assessments can provide more credible evidence of causation than randomized experiments or longitudinal studies (Kaplan, 2009; Schneider, Carnoy, Kilpatrick, Schmidt & Shavelson, 2007). Different study designs have advantages and disadvantages for establishing causation and there is no gold standard. For example, random assignment to treatment and control conditions in randomized experiments yields evidence of causal effects, but randomized experiments suffer from the same problems as international assessments for producing evidence of causal mechanisms, because the mediating variable is not randomized, only the treatment is. Also, international assessments typically yield higher external validity than randomized experiments and longitudinal studies based on small samples and affected by attrition. And since causal evidence is always local, international assessments contribute very rich evidence from a variety of contexts, educational systems, and nations. Instead, it is not clear that the

mechanisms and effects observed in randomized and longitudinal studies replicate more broadly, beyond their local context (Deaton, 2010). In particular, selection bias and small sample sizes in randomized experiments cast doubts on the external validity of results. Further, research questions not possible to investigate with an experiment for ethical reasons can be addressed with international assessment data. For example, with appropriate techniques (e.g., propensity score matching) it is possible to compare performance of students in private and public schools controlling for the socio-economic characteristics of the student intake without actually having to assign students to different type of schools (Dronkers & Avram, 2009).

Longitudinal studies typically produce stronger evidence of causation than international assessment, because they allow us to control for prior academic achievement and examine changes in outcomes and treatments. However, international assessments are equally effective for evaluating the effect of time-invariant treatments (Wunsch, Russo, & Mouchart, 2010). Also, prospective longitudinal data is not always essential to identify the causal order of events. Sometimes retrospective responses or ordinary logic reveals the causal order of events even with cross-sectional designs. This study provides a useful example.

Purpose and contribution

This study explores the value of the causal mediation framework (Imai, Keele, Tingley, & Yamamoto, 2011; Imai & Yamamoto, 2013; Muthén, 2011) for investigating *causal mechanisms* in educational research using observational data. Causal mediation is illustrated using PIRLS 2011 data to examine the causal mediating role of early literacy activities in parental educational influences on reading performance. Causal mediation has already been applied in political science and psychology research (Imai, Keele, Tingley, & Yamamoto, 2011; Imai & Yamamoto, 2013; Muthén, 2011). This study extends causal mediation analysis to educational research.

The study design includes retrospective responses from parents on early literacy activities, before the child entered school, and information on parental education and reading performance when the student was in the fourth grade of school. It is assumed that early literacy activities precede reading performance in chronological order and that parental education is relatively time-invariant during this period (i.e., right before the child entered school and the fourth grade) and precedes early literacy activities. Thus, the longitudinal design with retrospective data complies with mediation assumptions (Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001) that the treatment (i.e., parental education) precedes the mediator (i.e., early literacy activities), and the mediator precedes the outcome (i.e., reading performance). The retrospective longitudinal design provides an example that research questions of a longitudinal nature can be addressed with the cross-sectional data of international assessments. But certainly, a number of caveats and limitations should be kept in mind, for example, biases arising from making judgements about the past in the present.

The observational data makes it possible to conduct research that would not be ethically permissible with an experimental design (e.g., assign children to families that promote early literacy activities and to families that do not). The national representative data yields significant external validity and international comparative mediation evidence in an area where most published studies use data from countries in North America or Europe. PIRLS includes data from a great variety of educational systems, socioeconomic contexts, and cultures, such as Indonesia, Colombia, Morocco, Botswana, Germany, and the United States. With that, it contributes more *democratic* evidence to the knowledge base of the causal mechanisms underlying parental education influences on student performance.

Causal mediation framework

The causal mediation framework was introduced and implemented in the R programming language by Kosuke Imai and colleagues (Imai, Keele, & Tingley, 2010; Imai,

Keele, Tingley, & Yamamoto, 2010, see also Pearl, 2001 and 2012). Methodologically, causal mediation analysis bridges traditional mediation analysis (Baron & Kenny, 1986; MacKinnon, 2008) with the potential outcomes framework (Holland, 1986; Rubin, 1974) to establish benchmarks for interpreting mediation results in causal terms.

The potential outcomes framework

The potential outcomes framework is used to formalize the concept of causation (Holland, 1986; Rubin, 1974). The framework consists of three elements, the treatment, T , the unit, i , and the potential outcome, $Y_i(t)$. Here, the outcome is reading performance (Y), the unit of interest (i) is the student, and the treatment (T) is parental education. Consider, for simplicity, $T=1$ represents assignment to a family where parents have attained university education levels and $T=0$, the control condition, assignment to a family with no university degrees. Under this framework each student has two potential outcomes, one is academic achievement resulting from the treatment, $Y_i(1)$, and the other is academic achievement in the control condition, $Y_i(0)$. The difference between both outcomes, $Y_i(1) - Y_i(0)$, is the individual treatment or causal effect.

In practice, it is not possible to observe both outcomes for the same student, i . Even in a randomized experiment if the student was assigned to a family with university degrees (i.e., the treatment), the outcome under the treatment, $Y_i(1)$, is observed, and if the student was assigned to a family with no university degrees (i.e., control group), the outcome under the control condition, $Y_i(0)$, is observed, but never both of them. The observed outcome is $Y_i = Y_i(T_i)$ for every student, i . But there is always a *counterfactual*, or an unobserved condition, for example, the academic achievement in a family with no university degrees for a student who was assigned to the treatment condition.

While it is not possible to obtain causal evidence at the individual level, randomization allows researchers to produce evidence of average treatment effect (ATE),

$E[Y_i(1)-Y_i(0)]$. Also, if certain assumptions are met, econometric techniques such as, instrumental variables and regression discontinuity can attain some level of ATE. Traditional mediation analysis, however, has evolved apart from this framework.

Traditional mediation analysis

In traditional mediation analysis (Baron & Kenny, 1986; MacKinnon, 2008) a mediating variable is used to explain the association between a treatment and an outcome variable. Mediation analysis is conducted in three steps. First the outcome is regressed on the treatment, second the mediator is regressed on the treatment, and third the outcome is regressed on the treatment and the mediator. In order to claim mediation, the treatment should precede and be significantly related to the outcome and the mediator in the first two regressions. In the third regression the treatment and the mediator precede the outcome, the mediator is significantly related to the outcome, and the effect of the treatment has reduced or disappeared compared to the first regression. Indirect or mediation effects are captured by this reduction. Direct effects are treatment effects in the third regression. Figure 1 depicts the analytical model.

Figure 1. Mediation model

The mediator in the relationship between parental education and reading performance is early literacy activities. Following mediation conditions, it is assumed that early literacy activities precede reading performance and parental education precedes early literacy activities. Parental education is significantly related to reading performance and early literacy activities, and parental education effects on reading performance reduce once early literacy activities are included in the model.

Causal mediation analysis

The potential outcomes framework can be extended to define causal mediation effects (Imai, Keele, Tingley, & Yamamoto, 2011). M_i is used to denote the observed mediator (i.e., the level of early literacy activities) given the treatment T_i , (i.e., the level of parental education). The mediator has two potential outcomes for each student, $M_i(1)$ and $M_i(0)$, the level of early literacy activities for parents with and without university education, respectively. But only one outcome is observed, that is, $M_i = M_i(T_i)$. For example, if the parents completed university $M_i(1)$ is observed but not $M_i(0)$.

In the original framework potential outcomes depend only on the treatment. In the causal mediation framework potential outcomes are a function of the treatment and the mediator, as in mediation analysis. $Y_i(t, m)$ is used to denote the potential outcome and $Y_i(T_i, M_i(T_i))$, the observed outcome given the treatment and the mediator. Mediation equations in the causal framework are represented by

$$Y_i(T_i, M_i(T_i)) = \alpha_1 + \beta_1 T_i + \xi_1^T X_i + \varepsilon_{i1} \quad \dots(1)$$

$$M_i(T_i) = \alpha_2 + \beta_2 T_i + \xi_2^T X_i + \varepsilon_{i2} \quad \dots(2)$$

$$Y_i(T_i, M_i(T_i)) = \alpha_3 + \beta_3 T_i + \gamma M_i + \xi_3^T X_i + \varepsilon_{i3} \quad \dots(3)$$

Total treatment effects are captured by $\hat{\beta}_1$, average direct effects (ADE) by $\hat{\beta}_3$, and mediation effects by $\hat{\beta}_2\hat{\gamma}$ or $\hat{\beta}_1 - \hat{\beta}_3$ in traditional mediation analysis. Causal mediation effects are defined by $Y_i(t, M_i(1)) - Y_i(t, M_i(0))$. By holding the treatment constant, indirect effects of the treatment through the mediator are isolated and all other causal mechanisms are eliminated. Indirect effects or causal mediation effects capture changes in reading performance that would occur if early literacy activities change from the value observed in the control condition (i.e., no university) to the value observed under the treatment condition (i.e., completed university), while holding the treatment constant. If parental education is unrelated to early literacy activities, then $M_i(1) = M_i(0)$ and the causal mediation effect is zero. As before, causal mediation effects at the individual level cannot be identified and the goal is to estimate average causal mediation effects (ACME), that is, $E[Y_i(t, M_i(1)) - Y_i(t, M_i(0))]$. In order to interpret $\hat{\beta}_2\hat{\gamma}$ as a valid estimate of ACME an additional assumption is required.

Sequential ignorability assumption

The sequential ignorability assumption consists of two conditions (Imai, Keele, Tingley, & Yamamoto, 2011): (1) the treatment is randomized and (2) the mediator is randomized. Two ignorability assumptions are made sequentially. First, that the treatment is ignorable given observed confounders and second, that the mediator is ignorable given the treatment and confounders. The first condition is satisfied in randomized experiments. Empirical analyses of observational data include a number of control variables related to the treatment and the outcome or employ matching techniques, for example, to try to satisfy this assumption. The second condition is not warranted even in randomized experiments because only the treatment is randomized, not the mediator. Analyses control for variables related to the mediator and the outcome to try to comply with this condition. Control variables are represented by X_i in Equations 1 to 3.

In this example, the first assumption implies that students are randomly assigned to parents with different educational levels. Or that no systematic differences between students with different parental educational levels exist once confounders are controlled. The second condition presupposes that different levels of early literacy activities are assigned randomly to students with the same characteristics and treatment status, that is, the same levels of parental education. These conditions, especially the ignorability of the mediator, are very strong and cannot be verified with observational data, since unobserved confounders can affect the mediator and outcome variables. Sensitivity analysis is employed to evaluate the sensitivity of results to the existence of such confounders.

Sensitivity analysis

Sensitivity analysis probes the sequential ignorability assumption or evaluates the validity of findings under potential violation of the sequential ignorability assumption (Imai, Keele, Tingley, & Yamamoto, 2011). It provides benchmarks for interpreting mediation results causally. Whenever the assumption is not rejected, estimated mediation effects can be interpreted as ACME. If mediation results are sensitive, they can change substantially when the sequential ignorability assumption is violated. The sensitivity parameter, ρ , measures the correlation between the error in the mediation model, ε_{i2} , and the error in the outcome model, ε_{i3} (see Equations 2 and 3). This correlation arises if unobserved confounders affect both mediator and outcome variables, because these variables are part of the two error terms. The sequential ignorability assumption implies that ρ equals zero. Interpretation of $\hat{\beta}_2\hat{\gamma}$ as ACME assumes that ρ equals zero. Sensitivity analysis evaluates the extent to which $\hat{\beta}_2\hat{\gamma}$ can be interpreted causally.

Empirical illustration: Causal mediation of early literacy activities

Literature review

Parental education and family SES are related to student performance in a number of countries (see for example with PIRLS, [citation removed for blind review]). In part, this association is explained by the home literacy environment (HLE), that is, the environment the family provides the child to stimulate precursors of reading and spelling, and linguistic competencies (e.g., Farkas & Beron, 2004). Early literacy activities such as shared book reading (e.g., Bradley, Corwyn, Burchinal, McAdoo, & Garcia-Coll, 2001), storytelling (e.g., Snow & Dickinson, 1990; Watson, 2002), learning the alphabet, numbers, and letters (e.g., Parker, Boak, Griffin, Ripple, & Peay, 1999), playing rhyming games (e.g., Baker, Serpell, & Sonnenschein, 1995), and access to cognitively stimulating toys and books (e.g., Gottfried, Fleming, & Gottfried, 1998) are a key element of the home literacy environment that sets the stage for children's literacy development. And the possibilities for promoting early literacy activities within the family are related to the SES and educational levels of parents.

Although there is variability in the quality of the home learning environment both across and within all SES levels, there is evidence that, in general, children in higher income homes are more likely to experience a more positive learning environment than children from low income homes (e.g., Niklas & Schneider, 2010). Children in low SES households have less exposure to books at home and are less likely to be regularly read to by parents (e.g., Evans, 2004). Also, there is evidence that parents with higher educational levels use more developmentally enhancing activities, interact more verbally with the child, read to the child more often, and encourage book-sharing (Gustafsson, Yang-Hansen, & Rosén, 2013; Raikes et al., 2006). Early literacy activities and the learning environment appear to mediate almost half of the SES gap in reading achievement (Duncan and Brooks-Gunn, 2000; Kiernan & Mensah, 2011; Myrberg & Rosen, 2009).

Longitudinal studies in the US suggest that early literacy activities mediate family SES effects on academic achievement. For example, Guo and Harris (2000) showed that cognitive stimulation in the home and, to a lesser extent, parenting style and the physical environment of the home are mediating factors affected by low income and influencing children's intellectual development. Votruba-Drzal (2006) showed that the quality of home environment during early and middle childhood explained a portion of the effects of income on academic skills. Aikens and Barbarin (2008) explained that the relation between SES and children's reading competence status is mediated by the HLE, but the HLE did not account for many of the SES differences in the rate at which children gained reading skills.

Also, in the UK Sylva et al., (2008) showed direct and indirect effects of HLE. Even after taking account of general cognitive ability those with a more favourable HLE had better outcomes. Using data from the UK Millennium Cohort Study, Kiernan & Mensah (2011) found positive parenting to mediate the effects of poverty and disadvantage on school achievement. In Germany, Niklas & Schneider (2013) found that HLE acts as a mediator between SES and reading and writing precursors.

Most HLE mediation studies have used data from a single developed country. Only a few studies have investigated the mediating role of HLE across countries (Myrberg & Rosén 2008; Park 2008; Gustafsson, Hansen & Rosén 2013). Park (2008) examined the mediating role of HLE in 25 countries using PIRLS 2001 data. He found that home literacy environments mediated the effect of parental education on reading achievement, but the mediating role of early home literacy activities was weaker than that of parental reading attitudes and the number of books. HLE mediating effects varied between 20% and 30% for most countries. The mediating role of early literacy activities was smaller, on average 10%. In a recent study with PIRLS 2011 data, Gustafsson, Hansen and Rosén (2013) also provide evidence of partial mediation of HLE by means of structural equation models.

Data

The data are sourced from PIRLS 2011. PIRLS 2011 is the most recent international assessment of reading performance (Y) in elementary school, when the study of early literacy skills seems more relevant. It included 53 education systems in very diverse contexts (e.g., Australia, Canada, Colombia, Croatia, Germany, France, Indonesia, Qatar, Kuwait, Sweden, United Arab Emirates). Parents reported early literacy activities (M) before the child entered school on a Likert-type scale: how often they read books, tell stories, sing songs, play with alphabet toys, talk about what you have read, play word games, write letters or words, and read aloud signs and labels with the child (1=never or almost never, 2= sometimes, 3=often). Responses to these 8 items were summarised into a single score using a partial credit model with item response theory (IRT). Parental education (T) is the maximum level of education of either parent (1=some primary, lower-secondary or no school, 2=finished lower-secondary, 3=finished upper-secondary, 4=finished post-secondary but not university, and 5=finished university or higher).

Models control for a number of confounders (X), including mother's and father's occupational status (see [citation removed for blind review] for calculation details), number of books at home, number of children books at home, a home possessions index (i.e., a Rasch score summarizing the possession of 5 items at home: a computer, a study desk, books, an own room, and internet connection), pre-school education status (i.e., attended pre-school or not), student's sex (i.e., girls and boys), language the student speaks at home (i.e., language of the test versus another language), frequency the student speaks the language of the test at home (i.e., never, sometimes, always), and the parents like reading scale (see Martin & Mullis [2012] for technical details).

Analyses

Analyses were conducted in R. Package [citation removed for blind review] was used for data management ([citation removed for blind review]), package *eRm* for creating Rasch scores for the mediator of early literacy activities and the home possessions index (Mair, Hatzinger, & Maier, 2012), package *Amelia* to carry out multiple imputation by country using the expectation-maximization (EM) algorithm (Honaker, King, & Blackwell, 2011), and package *mediation* to perform mediation and sensitivity analysis (Tingley, Yamamoto, Keele, and Imai, 2013). Five imputed datasets were created with *Amelia*. Analyses, however, were conducted on a single imputed dataset. The data has a multilevel structured but we conducted linear regression models because currently the mediation package cannot produce sensitivity analysis for multilevel models. That is, standard errors are underestimated: imputation uncertainty and serial correlation within schools are not taken into account.

Results

Figure 2 presents summary results of mediation analysis by education systems. It depicts three key parameters and corresponding confidence intervals (95%): ACME - the average causal mediation effects ($\hat{\beta}_2\hat{\gamma}$), ADE - the average direct effects ($\hat{\beta}_3$), and the total effects of parental education ($\hat{\beta}_1$). The data underpinning Figure 2 is reported in Table A1 (see Appendix A). ACMEs are significant in 18 education systems out of 50 (e.g., Bulgaria, Alberta [Canada], Saudi Arabia, Abu Dhabi, Chinese Taipei, Hong Kong, South Africa, Sweden, Trinidad and Tobago). In general mediation effects are small in comparison to total effects. In part, this is because the association between parental education and early literacy activities is not always positive. Table A2 (see Appendix A) records the coefficient of parental education (PARED) in a regression of early literacy activities (ELA) on PARED while controlling for confounders (i.e., $\hat{\beta}_2$ in Equation 2). The coefficient is positive and statistically significant in 25 education systems, half of the sample exactly.

In addition to the ACME estimates, Table A1 reports the percentage of the total effect of parental education mediated by early literacy skills. The maximum percentage is 10% in Bulgaria and the average percentage in education systems where ACME is significant is around 4%. In most education systems the total effect of parental education reflects the direct effect of parental education (see Figure 2). The total effects and direct effects of parental education are significant in all education systems except for Denmark. On average, the total effects after controlling for confounders ($\hat{\beta}_1$) amount to 10% of a standard deviation of the reading performance (SD= 100 score points).

Figure 3 presents sensitivity analysis for the ACME with respect to the correlation of the error terms in the mediation and outcome models, that is, the sensitivity parameter (ρ). The dashed line represents the estimated ACME when the sequential ignorability assumption is made (i.e., $\rho=0$). This is the ACME reported in Figure 2. The shaded region indicates confidence intervals (95%) for each value of ρ . Figure 2 indicated very small ACMEs. Further, results of Figure 3 indicated that statistically significant ACMEs are very sensitive to violations of the sequential ignorability assumption. A very small correlation between the error terms causes the ACME to become zero. In Bulgaria, for example, the country with the greatest ACME, a correlation smaller than 0.10 causes the ACME to be zero. The results are therefore quite sensitive to the presence of unobserved confounders related to early literacy activities and reading performance. For example, family cultural and economic capital variables not in the dataset could prevent us from interpreting the ACME in causal terms.

Although not essential for this paper, Table A2 (see Appendix A) reports the direct effect of ELA on reading performance once the treatment and confounders are controlled (i.e., $\hat{\gamma}$ in Equation 3). The results indicated a direct positive effect of ELA in most education systems (i.e., 41). That is, despite weak evidence for ELA being the causal mechanism

explaining parental education effects on reading performance, ELA did relate to reading performance directly independently of parental education levels.

Figure 2. Mediation results of parental education effects on reading performance

Figure 3. Sensitivity analysis for the early literacy activities mechanism

Discussion

Identifying evidence of causation is a fundamental goal of educational research and policy. Researchers are interested in the mechanisms whereby students, families, teachers, and governments influence educational outcomes. Policymakers seek for causal evidence to demonstrate the effectiveness of interventions in a context where resources are scarce and interventions can place a great strain on national budgets (e.g., hiring more teachers in order to reduce class size). The idea of using scientific evidence for claiming causation and informing policy decisions in education is not new. But in recent years educationalists have become particularly interested in statistical techniques for determining causal effects with observational data (e.g., regression discontinuity, propensity score matching, and instrumental variables). This paper has cautioned against overstated enthusiasm in their potential for driving the causal quest in educational research. It is argued that these techniques often oversimplify the reality of education systems involving complex interactions and multiple actors, because models concentrate on the estimation of treatment effects and rely strongly on exogenous variation not necessarily informed by theory. And that mediation techniques adapt better to the complex nature of education settings but until recently could not be applied under the causal framework. Also, it is argued that observational data from international assessments can contribute valuable evidence of causation. In light of these arguments, the paper has presented an application of causal mediation analysis with international assessment data.

The applied example examined the mediating role of early literacy activities in parental education effects on reading performance using retrospective responses in the data. Overall, the results provided weak support for the early literacy activities causal mechanism. Mediation effects were significant in 18 of the 50 education systems in the sample and the proportion of parental education effects mediated by early literacy activities was 10% or less.

Further, mediation effects were very sensitive to departures from the sequential ignorability assumption. That is, unobserved confounders related to early literacy activities and reading performance prevent us from interpreting mediation results in causal terms.

The results, however, were not very different from previous studies. For instance, cultural capital has been found to mediate 15% of family wealth effects on math achievement in the US (Orr, 2003) and early literacy activities mediated on average 10% of parental education effects on reading achievement in PIRLS 2001 (Park, 2008). But a direct comparison with previous studies is not possible due to differences in mediating variables (e.g., home learning environment, cultural capital), treatment variables (e.g., SES), control variables, educational outcomes, school stages, country samples, etc. For example, related studies have also found that parenting (i.e., activities to promote learning, interactions with the child, family organization and nutrition) mediates 40% of family resources influences in the UK (Kiernan & Mensah, 2011), that the number of books at home, early literacy activities, and the early reading abilities mediate on average 50% of parental education influences in Sweden (Myrberg & Rosén, 2009), and that home cognitive stimulation (e.g., learning materials, academic stimulation, physical environment) mediates 38% of maternal education influences in the US (Raviv, Kessenich, & Morrison, 2004).

In accordance with previous studies, the analyses provided support for direct effects of early literacy activities (Melhuish et al., 2008; Merlo, Bowman, & Barnett, 2007; Niklas & Schneider, 2013). These effects appeared to be independent of the family socio-economic background of students and more robust than indirect effects through parental education and family SES. Unreported analyses also showed that causal mediation of early literacy activities on family SES effects was very similar to mediation results for parental education. Parental education was preferred over family SES because the early literacy activities

mechanism was more direct for parental education than for family SES, which is a broader concept with no strong theoretical basis.

The analyses have some limitations. One is that retrospective responses on early literacy activities could be biased by present knowledge. Another is cross-cultural bias in the early literacy activities scale. Also, a broader concept such as the home learning environment would have facilitated comparisons with previous studies. But early literacy activities were preferred because responses were retrospective and consistent with mediation assumptions. Another limitation is that models do not take into account the missing data uncertainty and the multilevel nature of the data. However, the analyses are primarily intended to provide an illustration of causal mediation and results should be interpreted with caution. And yet another is that we have not explained the significant cross-national variation of treatment and mediation effects. Further analyses could examine whether these effects are related to cultural, social and economic characteristics of countries (e.g., individualism/collectivism, national income, welfare regime, stratification of school systems). For example, it could be that early literacy skills matter more in developed countries because developing countries still have to meet basic needs of children (e.g., nutrition, material resources; Park, 2008).

Further research on causation with international assessments

Educational science can benefit from methodological advances taking places in the areas of economics, psychology, and political science. The combination of methodological approaches holds promises for the study of causation. The econometric techniques alone are limited for addressing causal questions in educational settings. But they can be used in combination with mediation analysis for explaining causal mechanisms. For example, the instrumental variables approach can help identify exogenous sources of variation in the mediators. Together, these approaches can enhance the explanatory power of educational models.

Further methodological developments should also consider the presence of multiple mediators. This is important because the literature often points to different factors and theories explaining treatment effects. For example, studies on educational inequality suggest that family SES influences on student performance are mediated by pre-school education, paternal warmth and responsiveness, physical environment, home literacy environment, among other factors. And researchers are interested in the relative importance of these explanations and in evaluating competing theories. Statistical techniques for evaluating the presence of multiple mediators would therefore strengthen the methodological base for studying causation.

Despite its limitations, international assessment data offers many opportunities for addressing causal questions in educational research. A creative study design, guided by theory and using appropriate statistical techniques can produce very rich and persuasive evidence of causation. International assessments challenge researchers to find creative ways to take advantage of the great amount of variability available in the data. For example, researchers can combine international assessments with national datasets including prior achievement measures or longitudinal data. Increasingly, researchers can rely on trend data from international assessments (e.g., 12 years in PISA 2000-2012 and 16 years in TIMSS 1995-2011) to study change in outcomes and the reasons behind it. Researchers should also be mindful of exogenous sources of variation that create artificial experiments in the data, such as the introduction of policies between survey cycles (e.g., the establishment of the European Union) or from primary to secondary schooling (see, for example, school tracking effects in Hanushek and Wößmann, 2006) as well as rules for promotion, maximum class sizes (Angrist & Lavy, 1999), and age-of-school entry policies (Martin, Mullis, and Foy, 2011).

When defining a methodological framework for studying educational outcomes, educationalists should carefully consider the techniques and theories that adapt best to the nature of educational settings and the causal questions of relevance in educational research. Rather than adopting a unique view of causation, educational research can benefit from a multi-disciplinary approach that combines methods and theories from psychology, sociology, and economics. The specific statistical techniques are relevant but probably the least important aspect of a strong research study. More relevant aspects, for example, are that the study is guided by solid theory, the measures are valid, the sample yields some level of external validity, the design is appropriate for the research question, and competing mechanisms can be ruled out.

References

- Angist, J. D., & Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2), 533-575.
- Angrist, J.D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Artelt, C. (2005). Cross-cultural approaches to measuring motivation. *Educational Assessment*, 10(3), 231–255.
- Baker, L., Serpell, R., & Sonnenschein, S. (1995). Opportunities for literacy learning in the homes of urban preschoolers. In L. Morrow (Ed.), *Family literacy: Connections in schools and communities* (pp. 236-252). Newark: International Reading Association.
- Baird, J.-A., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T., & Daugherty, R. (2011). *Policy effects of PISA*. Oxford: Oxford University Centre for Educational Assessment.
- Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & Coll, C.G. (2001). The home environments of children in the United States Part II: Relations with behavioral development through age thirteen. *Child Development*, 72, 1868–1886.
- Farkas, G. & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, 33, 464–497.
- Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1998). Role of cognitively stimulating home environment in children's academic intrinsic motivation: A longitudinal study. *Child Development*, 69(5), 1448-1460.
- Deaton, A. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, 48, 424–455.

- Dronkers, J. & S. Avram, (2009). Choice and Effectiveness of Private and Public Schools in Seven Countries. A reanalysis of three Pisa Data Sets. *Zeitschrift für Pädagogik*, 55, 895-909.
- Duncan, G. J., & Brooks-Gunn, J. (2000). Family poverty, welfare reform, and child development. *Child Development*, 71, 188–196.
- Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology*, 54, 65–75.
- Grek, S. (2009). Governing by numbers: The PISA “effect” in Europe. *Journal of Education Policy*, 24, 23–37.
- Guo, G. & Harris, K.M. 2000 ‘The Mechanisms Mediating the Effects of Poverty on Children’s Intellectual Development’, *Demography*, 37(4), 431–47.
- Gustafsson, J-E. (2013). Causal inference in educational effectiveness research: a comparison of three methods to investigate effects of homework on student achievement. *School Effectiveness and School Improvement*, 24(3), 275-295.
- Gustafsson, J., Hansen, K., & Rosén, M. (2013). Effects of Home Background on student Achievement in Reading, Mathematics and Science at the Fourth Grade. In Martin, M.O. & Mullis, I.V.S. (Eds.) *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—Implications for early learning*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Hanushek, E. A., & Wößmann, L. (2006). Does educational tracking affect performance and inequality? Differences- in-Differences evidence across countries. *The Economic Journal*, 116 (510), 63–76.
- Heckman, J. J. (2008). Econometric causality. *International Statistical Review*, 76(1), 1–27.

- Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81, 945–60.
- Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, 45(7), 1–47.
- Hopfenbeck, T.N. & Maul, A. (2011) Examining Evidence for the Validity of PISA Learning Strategy Scales Based on Student Response Processes. *International Journal of Testing*, 11(2), 95-121.
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309 –334.
- Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2010). Causal Mediation Analysis Using R. In *Advances in Social Science Research Using R*, ed. H. D. Vinod, Lecture Notes in Statistics. Springer-Verlag: New York, 129–54.
- Imai, K., & Yamamoto, T. (2013). Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments . *Political Analysis*, 21(2), 1-31.
- Kaplan, D. (2009). Causal inference in non-experimental educational policy research. In G. Sykes, B. Schneider., & D. N. Plank (Eds.), *Handbook on Education Policy Research*. New York: Taylor and Francis.
- Kiernan, K. E. & Mensah, F. K. (2011) Poverty, family resources and children’s early educational attainment: the mediating role of parenting. *British Educational Research Journal*, 37(2), 317–336.
- Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How Do Risk Factors Work Together? Mediators, Moderators, and Independent, Overlapping, and Proxy Risk Factors. *The American Journal of Psychiatry*, 158(6), 848-856.

- MacKinnon, D. P. (2008). Introduction to statistical mediation analysis. Mahwah, NJ: Erlbaum.
- Mair, P., Hatzinger, R., & Maier, M. J. (2012). eRm: Extended Rasch Modeling. R package version 0.15-1. <http://CRAN.R-project.org/package=eRm>
- Marsh, H. W., Hau, K-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, 6, 311–360.
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2011). Age distribution and reading achievement configurations among fourth-grade students in PIRLS 2006. *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, 4, 9-33.
- Martin, M.O. & Mullis, I.V.S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Merlo, L. J., Bowman, M., & Barnett, D. (2007). Parental nurturance promotes reading acquisition in low socioeconomic status children. *Early Education and Development*, 18, 51–69.
- Murnane, R. J., & Willett, J.B. (2011). *Methods matter. Improving causal inference in educational and social research*. New York: Oxford University Press.
- Muthén, B. (2011). *Applications of causally defined direct and indirect effects in mediation analysis using SEM in Mplus*. Manuscript submitted for publication.
- Myrberg, E. & Rosén, M. 2008, A path model with mediating factors of parents' education on students' reading achievement in seven countries. *Educational Research and Evaluation*, 14(6), 507-520.

- Myrberg, E & Rosen, M (2009). Direct and indirect effects of parents' education on reading achievement among third graders in Sweden. *British Journal of Educational Psychology*, 79, 695–711.
- Niklas, F & Schneider, W (2013). Home Literacy Environment and the beginning of reading and spelling. *Contemporary Educational Psychology*, 38, 40–50.
- Orr, A. J. (2003). Black-white differences in achievement: The importance of wealth. *Sociology of Education*, 76, 281–304.
- Park, H. (2008). Home Literacy Environments and Children's Reading Performance: A Comparative Study of 25 Countries. *Educational Research and Evaluation*, 14, 6, 489-505.
- Parker, F. L., Boak, A. Y., Griffin, K. W., Ripple, C., & Peay, L. (1999). Parent–child relationship, home learning environment, and school readiness. *School Psychology Review*, 28, 413–425.
- Pearl, J. (2001). Direct and indirect effects. In J. Breese & D. Koller (Eds.), *Uncertainty in artificial intelligence, proceedings of the seventeenth conference* (pp. 411–420). San Francisco: Morgan Kaufmann.
- Pearl, J. (2012). The Causal Mediation Formula—A Guide to the Assessment of Pathways and Mechanisms. *Prevention Science*, 13(4):426-36.
- Raikes, H., Pan, B. A., Luze, G., Tamis-LeMonda, C. S., Brooks-Gunn, J., Constantine, J., et al. (2006). Mother–child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development*, 77(4), 924–953.
- Raviv, T., Kessenich, M., & Morrison, F. J., (2004). A mediational model of the association between socioeconomic status and three-year-old language abilities: The role of parenting factors. *Early Childhood Research Quarterly*, 19, 528–547.

- Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies. *Journal of Educational Psychology*, 66, 688–701.
- Schlotter, M., Schwerdt, G., & Woessmann (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics*, 19(2), 109-137.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimation causal effects: Using experimental and observational designs*. Washington, DC: American Educational Research Association.
- Snow, C. E., & Dickinson, D. K. (1990). Social sources of narrative skills at home and at school. *First Language*, 10, 87–103.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). *Effective Pre-School and Primary Education 3–11 project (EPPE 3–11): Final report from the primary phase: Pre-school, school and family influences on children’s development during Key Stage 2 (age 7–11)*. London: University of London.
- Tingley, D., Yamamoto, T., Keele, L., & Imai, K. (2013). *mediation: R Package for Causal Mediation Analysis. R package version 4.4*. URL <http://CRAN.R-project.org/package=mediation>
- Van Damme, J., Liu, H., Vanhee, L., & Pustjens, H. (2010). Longitudinal studies at the country level as a new approach to educational effectiveness: explaining change in reading achievement (PIRLS) by change in age, socio-economic status and class size. *Effective Education*, 2(1), 53-84.
- VanderWeele T.J. & Vansteelandt S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, 2, 457-468.
- Votruba-Drzal, E. (2006). Economic disparities in middle childhood development: Does income matter? *Developmental Psychology*, 42, 1154–1167.

- Watson, R. (2002). Literacy and oral language: Implications for early literacy acquisition. In S. B. Neuman & D.K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 43–53). New York: The Guilford Press.
- Wößmann, L. & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3), 695–736.
- Wunsch G., Russo F., & Mouchart M. (2010). Do we necessarily need longitudinal data to infer causal relations? *Bulletin de Méthodologie Sociologique*, 106(1), 1-14.

Appendix A

Table A1. Causal mediation estimates

Education system	% Mediated	ACME	Direct	Total
Andalusia, Spain	0.56	0.06	9.94 *	10.00 *
Abu Dhabi, UAE	5.06	0.77 *	14.50 *	15.27 *
Australia	0.06	0.01	10.69 *	10.70 *
Austria	0.41	0.03	6.52 *	6.55 *
Azerbaijan	0.00	0.00	5.04 *	5.01 *
Belgium (French)	0.96	0.07 *	6.61 *	6.69 *
Botswana	2.88	0.36 *	12.30 *	12.66 *
Bulgaria	10.32	1.10 *	9.57 *	10.66 *
Canada	0.52	0.05	10.04 *	10.09 *
Canada, Alberta	7.54	0.65 *	7.92 *	8.57 *
Canada, Ontario	0.00	0.00	10.63 *	10.45 *
Canada, Quebec	1.01	0.12	10.71 *	10.83 *
Chinese Taipei	3.67	0.42 *	10.60 *	11.02 *
Colombia	0.00	0.00	7.70 *	7.70 *
Croatia	0.00	0.00	6.22 *	6.21 *
Czech Republic	0.00	0.00	5.82 *	5.65 *
Denmark	3.32	0.06	1.47	1.53
Dubai,UAE	2.97	0.60 *	19.33 *	19.93 *
Finland	0.00	0.00	7.79 *	7.75 *
France	0.00	0.00	12.59 *	12.50 *
Georgia	0.34	0.06	13.28 *	13.33 *
Germany	0.00	0.00	6.35 *	6.32 *
Honduras	1.53	0.18 *	11.14 *	11.32 *
Hong Kong SAR	3.21	0.14	3.82 *	3.96 *
Hungary	0.00	0.00	11.55 *	11.29 *
Indonesia	0.09	0.01	13.44 *	13.45 *
Iran, Islamic Rep. of	0.09	0.02	13.34 *	13.35 *
Italy	0.00	0.00	8.31 *	8.26 *
Kuwait	0.07	0.02	18.77 *	18.79 *
Lithuania	0.76	0.06	6.92 *	6.98 *
Malta	5.12	0.70 *	12.42 *	13.12 *
Morocco	0.00	0.00	13.87 *	13.29 *
Netherlands	0.01	0.00	5.28 *	5.29 *
New Zealand	2.51	0.26 *	10.03 *	10.29 *
Northern Ireland	0.00	0.00	6.67 *	6.52 *
Norway	0.00	0.00	3.28 *	3.27 *
Oman	3.67	0.50 *	13.06 *	13.56 *
Poland	0.82	0.12	14.00 *	14.12 *
Portugal	0.03	0.00	4.83 *	4.83 *
Qatar	1.54	0.32 *	20.09 *	20.41 *
Romania	0.93	0.15	15.95 *	16.10 *
Russian Federation	0.55	0.06	9.94 *	10.01 *

Saudi Arabia	5.41	0.55 *	9.36 *	9.91 *
Singapore	1.92	0.29 *	14.11 *	14.40 *
Slovak Republic	0.00	0.00	7.41 *	7.38 *
South Africa	2.74	0.47 *	16.55 *	17.03 *
Spain	0.00	0.00	6.93 *	6.90 *
Sweden	4.63	0.35 *	7.07 *	7.41 *
Trinidad and Tobago	4.57	0.46 *	9.44 *	9.90 *
United Arab Emirates	4.10	0.70 *	16.24 *	16.93 *

Notes.

* $p < 0.05$.

Negative ACMEs were coded as 0.

ACME=Average causal mediation effect;% Mediated = percentage mediated of the total parental education effect; Direct = direct effect of parental education; Total = total effect of parental education.

Table A2. PARED coefficient in ELA model and ELA coefficient in READ model

Education systems	ELA model PARED			READ model ELA		
	Estimate	Std. Error		Estimate	Std. Error	
Andalusia, Spain	0.02	(0.02)		3.32	(0.80)	*
Abu Dhabi, UAE	0.07	(0.02)	*	10.92	(1.20)	*
Australia	0.00	(0.02)		4.44	(0.67)	*
Austria	0.02	(0.02)		1.61	(0.70)	*
Azerbaijan	0.06	(0.02)	*	-0.49	(0.80)	
Belgium (French)	0.04	(0.02)	*	1.76	(0.80)	*
Botswana	0.10	(0.02)	*	3.77	(0.90)	*
Bulgaria	0.28	(0.02)	*	3.82	(0.71)	*
Canada	0.01	(0.01)		3.95	(0.33)	*
Canada, Alberta	0.13	(0.03)	*	5.13	(0.81)	*
Canada, Ontario	-0.05	(0.03)	*	3.48	(0.78)	*
Canada, Quebec	0.03	(0.03)		3.74	(0.75)	*
Chinese Taipei	0.13	(0.02)	*	3.26	(0.80)	*
Colombia	0.04	(0.02)	*	-0.05	(0.86)	
Croatia	0.00	(0.02)		3.05	(0.71)	*
Czech Republic	-0.07	(0.02)	*	2.39	(0.74)	*
Denmark	0.03	(0.02)		1.85	(0.77)	*
Dubai,UAE	0.06	(0.02)	*	9.54	(0.96)	*
Finland	-0.02	(0.02)		2.67	(0.84)	*
France	-0.03	(0.02)		3.24	(0.75)	*
Georgia	0.02	(0.02)		2.30	(0.78)	*
Germany	-0.03	(0.02)		0.93	(0.77)	
Honduras	0.09	(0.02)	*	2.01	(0.80)	*
Hong Kong SAR	0.04	(0.02)	*	3.69	(0.86)	*
Hungary	-0.06	(0.02)	*	3.99	(0.76)	*
Indonesia	0.09	(0.02)	*	0.02	(0.90)	
Iran, Islamic Rep. of	0.04	(0.02)	*	0.39	(0.81)	
Italy	-0.01	(0.02)		3.88	(0.80)	*
Kuwait	0.00	(0.02)		4.50	(1.44)	*
Lithuania	0.01	(0.02)		4.25	(0.75)	*
Malta	0.08	(0.02)	*	8.80	(1.15)	*
Morocco	0.11	(0.02)	*	-5.22	(0.71)	*
Netherlands	0.00	(0.02)		0.69	(0.70)	
New Zealand	0.05	(0.02)	*	5.70	(0.75)	*
Northern Ireland	-0.04	(0.02)		4.36	(0.85)	*
Norway	0.00	(0.03)		3.76	(0.87)	*
Oman	0.07	(0.01)	*	7.33	(0.78)	*
Poland	0.01	(0.02)		8.37	(0.78)	*
Portugal	0.00	(0.02)		1.13	(0.83)	
Qatar	0.06	(0.02)	*	5.70	(1.21)	*

Romania	0.21	(0.03)	*	0.67	(0.80)	
Russian Federation	0.03	(0.03)		2.15	(0.73)	*
Saudi Arabia	0.09	(0.01)	*	5.94	(1.06)	*
Singapore	0.11	(0.02)	*	2.57	(0.70)	*
Slovak Republic	-0.02	(0.02)		1.70	(0.66)	*
South Africa	0.08	(0.02)	*	5.96	(1.19)	*
Spain	-0.01	(0.01)		4.96	(0.59)	*
Sweden	0.05	(0.02)	*	6.88	(0.73)	*
Trinidad and Tobago	0.04	(0.02)	*	10.98	(1.01)	*
United Arab Emirates	0.07	(0.01)	*	9.70	(0.63)	*

* $p < 0.05$