Beyond the Stars

March 2019

Abstract

It is frequent to hear in economic seminars or read in academic papers that an effect is economically significant or economically important. Yet, the economic literature is vague on what economic importance means and how it should be measured. In this paper, I show that existing measures of economic importance are flawed and misused. I derive a new metric which measures, in percentage terms, the contribution of each explanatory variable to deviations in the dependent variable, ceteris paribus. As an illustration, the method is applied to study the determinants of migration and the determinants of fertility.

Keywords: Economic importance, Effect size, Migration, Fertility
JEL Classification: B4, C18, O10, F22, J10

1 Introduction

Controversy erupted two decades ago when McCloskey and Ziliak (1996) highlighted that 70% of articles published in the American Economic Review in the 1980s did not distinguish between economic and statistical significance. One decade later, the same authors observed little change in this trend: 82% of papers published in the American Economic Review in the 1990s *"mistook a merely statistically significant finding for an economically significant finding"* (Ziliak and McCloskey 2004).[1]

The practice in economics has changed considerably since then. It is now common to hear in economic seminars or read in academic papers that an effect is "economically significant" or "economically important". This trend is, of course, more than welcome for the credibility of the field. However, the fact that most empirical papers conclude that their findings are economically important also raises questions. Such claims could indeed result from a flawed definition or method, or be a cheap way to oversell results and increase their marketability.

What do economists understand by economic importance? What properties should measures of economic importance satisfy? Do existing methods correspond to this common understanding while satisfying key properties? If not, is it possible to design a new measure fulfilling these objectives? This paper aims to answer these four questions, which have been overlooked in the literature.

First, I explore what economic importance means for economists, building on survey data collected from economists who participated in an online experiment. According to a majority respondents, the economic importance of a variable refers to the size of its contribution to deviations[2] in the level of the dependent variable.

Second, I critically review existing methods to assess the economic importance of variables. I show that commonly-used methods - such as standardized beta

[1]The emphasis on statistical significance at the expense of effect size is also a problem in other fields (Ziliak and McCloskey 2008), implying that the contribution of this paper extends beyond economics.

[2]I use the terminology "deviations" to refer to changes in the dependent variables. Deviations in a variable can for example be measured by the standard deviation, or by the mean absolute deviation from a central point (usually the mean or the median).

coefficients, the mean value decomposition of Holgersson et al. (2014), the Shapley value, or the partial r^2 - are flawed and challenging to interpret.

Third, I derive a new measure of economic importance of variables and residuals in regressions. For each variable, I derive an indicator of absolute economic importance that measures the proportion of deviations in the dependent variable explained by the explanatory variable under scrutiny, ceteris paribus. The new measure is expressed in percentage terms, making it particularly straightforward to interpret. The relative economic importance of two variables is given by the ratio of their absolute economic importance. Two limitations should be acknowledged. First, the new measure of importance focuses on ceteris paribus effects, that is, on the partial derivatives of the dependent variable with respect to explanatory variables. This means that the new measure of importance is usually uninformative about total derivatives. Second, the estimator of the new measure is consistent if regression coefficients are consistently estimated. But, as other measures of importance, it is biased in small samples (especially in the presence of multicollinearity), and biased and inconsistent in case of omitted variable bias, reverse causality bias, mismeasurement, or misspecification bias. Consequently, the new measure should not be considered as a magic bullet that can be applied blindly. Before assessing variables' importance, researchers should first solve or minimize endogeneity issues. Building on theory, researchers should also discuss whether partial and total derivatives are likely to differ, and how.

Finally, the new method is applied to three case studies. The first application is a monte-carlo experiment, which illustrates the intuitive appeal of the new metrics, while underling the weaknesses of existing methods.

The second application explores the determinants of international migration to OECD countries. The literature has identified the key determinants of skilled and unskilled migration (Docquier and Rapoport 2012): geographic and cultural distances, incomes in the origin and destination countries, restrictive or selective migration policies, and the presence of a large diaspora in destination countries. The relative economic importance of these different factors is however unclear.

I apply the new method to the dyadic data of Dao et al. (2018), and find that migration flows to OECD countries mainly depend on two factors: the size of the diaspora in destination countries (36.1%), and GDP per capita in origin countries (11.7%). This latter relationship is non linear: migration rates are increasing in GDP per capita below $2,762 per capita in PPP value, and decreasing in GDP per capita past this threshold. The increasing section of the inverted-U is mainly driven by the very low realization rates of less educated individuals living in poor countries. In contrast, I find that geographic and cultural barriers only play a minor role. Geographic distance explains 5.4% of deviations in migration rates, while sharing a common language and genetic distance are responsible for 4.8% and 3.2% of deviations respectively.

The third case study extends the work of de la Croix and Gobbi (2017) on the determinants of fertility. While the key factors influencing fertility are well-known (e.g. age, marital status, education, child mortality, population density), the literature is silent on their relative economic importance. The new method highlights that women's fertility is mainly explained by their age (46.6%), their level of education (10.2%) and their marital status (11%). Experiencing the death of a child is also associated with increased fertility (7.9%). Depending on the specification, population density explains between 3 and 10% of predicted deviations in fertility rates. While statistically significant, GDP per capita, child mortality at the cluster level, land quality, and distance to water only explain a tiny proportion of deviations in fertility. This example highlights the importance of assessing economic importance when analyzing large datasets, with which minor differences and unimportant factors can be statistically significant.

Why is this research important? In particular, why is it important to have a tool to compare the economic importance of different variables, while most empirical papers in economics focus on one mechanism at a time? The ultimate objective of the profession is not to understand the effect of a single factor on a dependent variable, but rather to map the full causal chain of economic phenomena. It is therefore crucial to have a generic method for measuring and comparing the economic importance of the multiple statistically significant causes of a phenomenon.

Having a credible measure of economic importance is all the more important given the increasing popularity of "Big Data", with which most tests are expected to be statistically significant.

Cost-benefit analysis is the first-best approach to measure economic importance. Cost-benefit analysis is widely used in fields where costs and benefits can be estimated (for example in health economics or in the impact evaluation literature). Rather than simply identifying which interventions produce a statistically significant effect, cost-benefit analysis identifies which intervention is the most cost-effective and should be expanded.

But for many research questions, cost-benefit analysis cannot be used because the costs or benefits do not exist or are too complex to be calculated. In the long-run growth literature, for example, it would be absurd to calculate the costs of modifying explanatory variables (e.g. the costs of reducing the slave trades or the cost of optimizing genetic diversity). Researchers contributing to the fertility or migration literature might be interested to understand the determinants of the phenomenon they study, but not be willing to calculate the social utility of having one more child or one more migrant. A method similar to cost-benefit analysis is needed to assess the economic importance of findings from descriptive studies that aim at understanding the causes of a phenomenon (e.g. long-run growth, fertility, migration), but that do not aim to, or cannot, directly affect this phenomenon. Just as it is important to compare the cost-effectiveness of different interventions using RCTs, it is also crucial to have a generic method to compare the economic importance of different variables in this type of descriptive study.

The contribution of this paper is to rigorously study "what economic importance means" and "how to measure it". It complements recent works on how to raise standards and increase transparency in applied economics (see e.g. Brodeur et al. (2016); Franco et al. (2014) on p-value hacking and publication bias; Maniadis et al. (2017) or Camerer et al. (2016) on replicability; Pritchett and Sandefur (2015) or Vivalt (2015) on external validity; Oster (2014) or Altonji et al. (2005) on the impact of unobservables; Miguel et al. (2014) on experimentation in social

sciences; Olken (2015) on pre-analysis plans).

The paper proceeds as follows. Section 2 clarifies the concepts of absolute and relative economic importance for economists. Section 3 scrutinizes existing methods to identify their strengths and flaws. Building on these observations, a new method for measuring absolute and relative economic importance is derived in Section 4. The new method is applied to study the determinants of migration and the determinants of fertility in Section 5. Section 6 concludes the study.

2 What does "economic importance" mean?

I propose that an empirical result can be deemed economically important if two conditions are met. First, the dependent variable under scrutiny must be economically relevant. Judging whether a variable is economically interesting is highly subjective, especially since economists enjoy exploring research questions that are relevant to other fields of study (chiefly psychology, history, demographics, and political science). Researchers should therefore rely on their own value judgment to determine whether this first condition is satisfied. Second, the effect of the explanatory variable on the dependent variable must be economically sizable. This second judgment is much more objective than the first one, as the size of an effect can be measured with an appropriate metric. This paper therefore focuses on this latter aspect, assuming that the dependent variable under scrutiny is economically relevant. Importantly, the objective assessment of whether an effect is sizable should not be confounded with the subjective assessment of whether the effect is unexpected or interesting.

The recent literature abounds with statements about the "economic significance" or "economic importance" of variables of interest. In all fields of economics, such statements are usually preceded or followed by an analysis of the predicted impact of an increase in the variable of interest on the dependent variable. The increase in the variable of interest is often expressed in standard deviation terms, in percentage terms, or as a jump from one quantile to another. The impact on the dependent variable is either expressed in level, in percentage of the mean, or

in standard deviation terms. A simple Google Scholar search for "economically important" generates the following examples. In a study of the determinants of ethnolinguistic diversity, Michalopoulos (2012) reports standardized coefficients to *"facilitate comparison of the quantitative effect across different specifications and across regressors."* Noting that *"a one-standard-deviation increase in the variation of elevation and a similar increase in the variation of land quality augments linguistic diversity by 0.31 and 0.34 standard deviations"*, the author concludes that *"these economically important findings reveal the geographic origins of contemporary ethnolinguistic diversity."* In their study of the effect of individual experiences of macroeconomic shocks on risk taking, Malmendier and Nagel (2011) conclude that *"the magnitudes of the effects are economically important"*, because *"an increase of 5 pp in the level of experienced real stock returns is associated with an increase in the probability of stock market participation of about 10 pp."* Also in finance, Djankov et al. (2008) argue that their *"six measures of the regulation of self-dealing are statistically and economically significant predictors of stock market development"* because *"the estimated coefficients imply that a two-standard-deviation increase in the indices of ex ante and ex post private control of self-dealing is associated with an increase in the ratio of market cap to GDP of 32 and 34 percentage points, respectively."* In their study of monetary policies, Aghion et al. (2009) report that the effect of exchange rate volatility on productivity growth *"is economically important: an increase of 50% in exchange rate volatility [...] leads to a 0.33% reduction in annual productivity growth."* The economic literature is replete with similar statements.

I make two observations from these quotations. First, the economic importance of an explanatory variable seems to refer to the size of its impact on deviations in the dependent variable. Contrary to the theoretical literature on variables' importance, which mostly focuses on variance decomposition,[3] the economic importance of an explanatory variable does not relate to the proportion of the variance or of the R^2 which is explained by this explanatory variable.

[3]See e.g. Grömping (2015), Johnson and LeBreton (2004), or Darlington (1968) for excellent reviews.

Second, assessing the economic importance of explanatory variables has the double objective of comparing the *magnitude* of the coefficients *across regressors* and *across different specifications*. In what follows, I therefore distinguish the concepts of *relative* and *absolute* economic importance. On the one hand, one may be interested in comparing the relative importance of different explanatory variables. For example, a researcher may wonder whether it is wealth or distance from a healthcare center which is the best predictor of health outcomes, and he or she may want to compare their *effect size*, that is, their relative importance. On the other hand, one may be interested to know the absolute economic importance of an explanatory variable. Using the same example, the researcher may be interested to know the proportion of changes in the dependent variable - here health outcomes - which is explained by a variable of interest, here wealth or the distance from a health care center.

To further confirm these observations, I ran an online experiment among 213 economists who are current or former colleagues from the University of Oxford, the University of Namur, or the Catholic University of Louvain, or who participated in the 2015 or 2016 CSAE conferences at the University of Oxford. Respondents were contacted by email.[4] Descriptive statistics are presented in Table A.2 in Appendix. This sample is of course not random, but I argue that it is quite representative of the population of economists I am interested in. For example, 60% of the sample have a PhD while 40% are PhD students. In comparison, the ratio of D.Phil. students to faculty is about 61% at the Department of Economics of the University of Oxford.[5] Those in the sample are highly-skilled in econometrics. When asked how often they are exposed to regression analysis in their daily work, on a scale from 1 "Never" to 10 "very regularly", the respondents replied with an average score of 7.7 (median=8, mode = 10) (see Figure A.4 in appendix A). In

[4]On the 24th of October 2016, I sent a personal email to all members of the Department of Economics at Oxford University (300 emails), and to all participants of the 2015 or 2016 CSAE conferences (480 emails). I organized a pilot survey in June 2016 with a selected sample of colleagues and former colleagues from the University of Oxford, the University of Namur (UNamur) and the Catholic University of Louvain (UCLouvain).

[5]On the 15th of March 2017, there were 83 D.Phil. students, 34 postdoctoral researchers, 51 associate professor, professors or readers, and 44 economists associated with Oxford colleges or other departments.

appendix B, I assess the robustness of results to changes in the composition of the sample. Responses are only marginally affected by respondents' proficiency in regression analysis or by having a PhD diploma. This suggests that the sample is quite representative of the population of interest despite the fact that it was not randomly drawn.

The online questionnaire included the following problem: *"Three normal variables x_1, x_2 and x_3 were randomly generated following a normal distribution N(0,1). These variables are uncorrelated. They are combined as follows to generate the variable y: $y = c_0 + c_1x_1 + c_2x_2 + c_3x_3$. We assume that y, x_1, x_2 and x_3 are variables that are of high interest for economists and policy-makers (e.g. welfare, wages, health, age, etc.). A linear regression ($R^2 = 1$) of y on x_1, x_2 and x_3 can be used to obtain the coefficients c_1, c_2 and c_3. In percentage, how would you qualify the economic importance of the effect of x_1, x_2 and x_3 on y?"*

I was expecting three possible answers to this question, each corresponding to a different understanding of the concept of absolute economic importance. The economic importance of x_i may refer to (i) its contribution to the variance of the dependent variable $c_i^2/({c_1}^2 + {c_2}^2 + {c_3}^2)$, (ii) its effect compared to the mean of the dependent variable c_i/c_0, and (iii) its contribution to deviations in the dependent variable $|c_i| / (|c_1| + |c_2| + |c_3|)$.

Two vectors of coefficients c_i were randomly assigned to participants.[6] Half of the sample was randomly assigned to the vector $(0, 1, 5, 2)$, which allows relatively easy computation for the interpretations (i) and (iii), but may be confusing for interpretation (ii). The other half was randomly assigned to the vector $(6, 1, 5, 2)$, which allows straightforward computation for the three interpretations. Table A.3 shows that respondents' characteristics are well balanced across randomization groups.

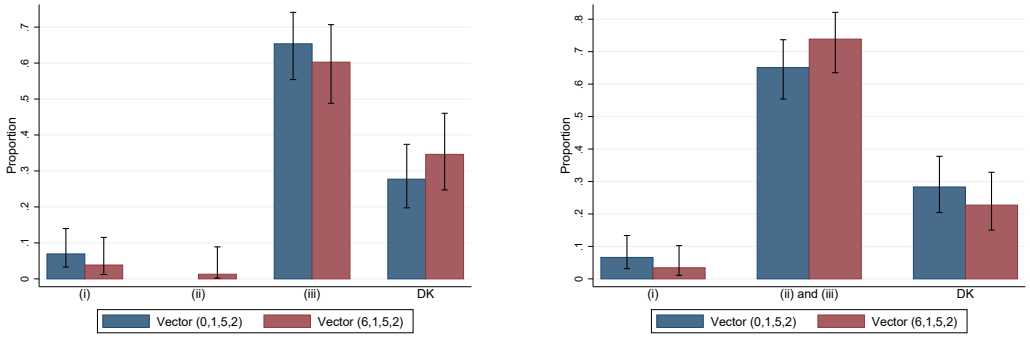
The results of this experiment are presented in Figure 1(a).[7] When facing the first vector, a large majority of respondents (65%) offered responses that are consis-

[6] I chose the coefficients such that the three interpretations require a similar level of effort for someone well-trained in econometrics.

[7] The categorization of responses is explained in appendix B.

tent with the interpretation (iii), confirming the hypothesis that most economists interpret the economic importance of a variable as its contribution to deviations in the dependent variable. Other respondents either did not know what to answer (28%), or had in mind the interpretation (i), on variables' contribution to the variance of the dependent variable (7%).

With the second vector $(6, 1, 5, 2)$, 60% of respondents offered responses corresponding interpretation (iii). Only 4% of responses corresponded to the interpretation (i), and one response corresponded to the interpretation (ii). As much as 35% of respondents replied that they could not reply to the question. As shown in appendix B, these results are not significantly correlated with respondents' exposure to regression analysis or with their qualification.



(a) Absolute importance

(b) Relative importance

Figure 1 – Results of experiment 1 ((i) to (iii) = response in line with interpretations (i) to (iii), DK = do not know).

The online survey also included a question about the relative economic importance of explanatory variables. Respondents were almost unanimous, and answers do not vary significantly across the two vectors of coefficients (Figure 1(b)). When asked to assess the relative economic importance of x_1 versus x_2, 69% of respondents answered 5/1, which corresponds to the ratio c_2/c_1. This answer is consistent with interpretations (ii) and (iii). Only 5% of respondents answered 25/1. This latter answer compares variables' contributions to the variance of the dependent variable, and therefore pertains to the interpretation (i). These results suggest that the relative economic importance refers to the ratio of contributions to de-

viations in the dependent variable, and not to the ratio of contributions to its variance.

In conclusion, a majority of economists understand the absolute economic importance of a variable as its contribution to deviations in the dependent variable. This interpretation makes sense when it comes to policy implications: policymakers are usually interested in influencing the level of the dependent variable, and not its variance or the r^2 of the relationship. The relative economic importance of two variables refers to the ratio of these contributions expressed in the same unit of measurement.

3 A review of existing methods

In this section, I review the properties of the main methods used by economists to evaluate the economic importance of explanatory variables in regression analysis. The objective is to identify their main strengths and flaws, to ensure that the new method will have the same advantages while not suffering from the same weaknesses.

3.1 Standardized regressions

Following the pioneering work of Wright (1921), standardizing the dependent and explanatory variables has been used for decades in regression analysis (see Darlington (1968) or Bring (1994) for critical reviews of the method). It is probably the most widely-used method to interpret the economic importance of explanatory variables in applied economics (see e.g. Spolaore and Wacziarg (2009), Nunn and Qian (2014), Alesina et al. (2015), Satyanath et al. (2017), Valencia Caicedo (2018)) but also in other fields (Luskin 1991; Gelman 2008; Nakagawa and Cuthill 2007). The coefficients of a *standardized regression* are commonly referred to as *standardized beta coefficients*.

Are the standardized beta coefficients useful to assess the economic importance of the explanatory variables x_i? Yes and no. On the one hand, the standardized

beta coefficient of an explanatory variable is a good measure of its ceteris paribus effect on the dependent variable. Therefore, the ratios of standardized beta coefficients allow comparison of the partial effects of explanatory variables *within regressions*, and, as such, are good measures of relative importance. On the other hand, the standardized beta coefficients are flawed measures of absolute economic importance because they cannot be easily compared *across regressions*. I identify two interrelated problems.

The first problem is that standardized beta coefficients are highly dependent on the coefficients of correlation between explanatory variables, which leads to incoherent results. Consider for example the simple data generating process $y = x_1 + x_2$ with $x_i \sim N(0, 1)$ for $i \in 1, 2$. The coefficient of correlation between x_1 and x_2 is denoted ρ_{12}. In this setup, the standardized beta coefficients of x_1 and x_2 are equal to $1/\sqrt{2}$ if $\rho_{12} = 0$, equal to $1/2$ if ρ_{12} approaches 1, and equal to $+\infty$ if ρ_{12} approaches -1. The fact that the importance of a variable is unbounded and can approach infinity is counter-intuitive. The fact that two variables from the same model can have an importance approaching infinity is even more troubling. And the fact that the importance of x_1 and x_2 increase when ρ_{12} decreases from 0 to any negative value seems illogical because the dispersion of y is actually decreasing at the same time. This counter-intuitive example illustrates that the standardized beta coefficient is a poor measure of absolute importance. In particular, the standardized beta coefficient of a variable x_i can be large either because its partial derivative is large or because the variance of y is small. Standardized beta coefficients are therefore difficult to interpret and compare across different specifications, especially when explanatory variables are highly correlated.

The second problem is that standardized beta coefficients do not add up to an easily interpretable number. Assuming a simple model $y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \epsilon$, the sum of the standardized beta coefficients is equal to:

$$S = \sum_{i=1}^{n} b_i^* = \frac{\sum_{i=1}^{n} \beta_i \sigma_i}{\sqrt{{\sigma_\epsilon}^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \beta_i \sigma_i \beta_j \sigma_j}}, \tag{1}$$

where the standardized beta coefficient of x_i is denoted b_i^*, the standard deviation of x_i is denoted σ_i, and the coefficient of correlation between x_i and x_j is denoted ρ_{ij}. The sum of standardized beta coefficients can be very high - much larger than 1 - if the error term is relatively small and if the number of explanatory variables n is large. The sum S is also larger if the correlation between explanatory variables is low or negative. The fact that the coefficients of a standardized regression do not add up to a "simple" number such as 1, 0, or the R^2 implies that standardized beta coefficients cannot be easily interpreted. In particular, they cannot be interpreted as the percentage. Researchers missing this property face the risk of overestimating the economic importance of their variables of interest. This risk is particularly important when the R^2 and the number of explanatory variables are large.

To assess whether this fear is well-founded, the online experiment included a module examining how economists interpret the results of a standardized regression. Results, presented in Appendix A, confirm that economists are often tempted to interpret the coefficients of standardized regression as percentages. As a corollary, I conclude that economists may actually be interested in having a method to assess the economic importance of variables that can be interpreted in percentage terms.

3.2 Mean value decomposition

Achen (1982) and Holgersson et al. (2014) propose using a mean value decomposition to assess the economic significance of variables. Following this method, the mean value of the dependent variable is decomposed into the contribution of each variable and of the intercept: the economic importance of x_i is defined as $\beta_i \frac{\bar{x_i}}{\bar{y}}$.

One nice property of this method is that the contributions of each variable and of the intercept add up to 1. Contributions can therefore be interpreted in percentage terms.

This method however has three major weaknesses. First, its conclusions can dramatically change when variables are rescaled. For example, the economic importance of temperature in a given model will vary if temperature is expressed in

degrees Celsius, Kelvin or Fahrenheit. Second, this method breaches the inclusion property, which states that the importance of a variable should be non null if its regression coefficient is non null. Finally, the method is inconclusive if the dependent variable is centered in 0. These problems warrant caution when using this method and interpreting its results.

3.3 Methods to decompose the R^2

Several methods aiming at decomposing the R^2 or the variance of the dependent variable have been developed by scholars from different fields (see Grömping (2015) or Kruskal and Majors (1989) for excellent reviews of existing methods and associated challenges). I focus here on the Shapley value and the partial r^2, as these statistics have recently been used by influential scholars to assess the importance of their results (see e.g. Henderson et al. (2018), Ashraf and Galor (2013), Arbath et al. (2018), Dustmann and Okatenko (2014), Manchin and Orazbayev (2018)).

Shapley values divide the R^2 in shares associated with each variable. To calculate Shapley values, variables are entered one by one in the regression and their marginal contribution to the R^2 are recorded. This process is repeated for all possible permutations of regressors. The Shapley value of a variable is the average of the marginal contributions to the R^2. This approach was independently proposed by Lindeman (1980) and Kruskal (1987), and then reinvented under various names (e.g. "hierarchical partitionning" by (Chevan and Sutherland 1991) or "dominance analysis" by (Budescu 1993)). Stufken (1992) and Lipovetsky and Conklin (2001) uncovered the link between this measure and the game-theoretic concept of the "Shapley value" (Shapley 1953), explaining why economists tend to use this name. The work of Shorrocks (2013) has generated a new interest in this method among economists (Dustmann and Okatenko 2014; Henderson et al. 2018; Manchin and Orazbayev 2018).

The main problem associated with this method is that irrelevant variables may have a large Shapley value if they are correlated with an important regressor. This is true even in large samples. In technical terms, the Shapley value violates the

exclusion criterion (Sterck 2018; Grömping 2015). In contrast, an important variable may have a relatively low Shapley value if it is correlated with irrelevant regressors. This metric is also problematic because of "both computational burden and lack of understanding what is really calculated" (Grömping et al. 2006): Shapley values are neither measuring the partial derivative nor the total derivative of the dependent variable with respect to explanatory variables. This metric is also biased and inconsistent in presence of multicollinearity. In summary, this computationally-demanding method should be avoided: while irrelevant variables may falsely look important, relevant ones may be deemed unimportant, implying that Shapley values are impossible to interpret.

The partial r^2 compares the variation of interest to the sum of the unexplained variation and the variation of interest. In other words, it measures the proportion of the unexplained variation that is explained by the addition of the variable of interest to the model. This method has sometimes been used by economists to assess the importance of variables or the strength of instruments in regression analysis (see e.g. Lipsey and Steuer (1961), Hart (1965), or Shea (1997)). It has recently been used by Ashraf and Galor (2013) and Arbath et al. (2018) in order to assess the importance of genetic diversity in explaining development and conflict.

There are three main problems associated with the partial r^2. First, this measure heavily depends on the correlation between explanatory variables. If explanatory variables are not orthogonal, their partial r^2 may be very low even if they may have large regression coefficients. As a consequence, this metric is also highly vulnerable to multicollinearity issues. Second, the sum of partial r^2 is not an easily interpretable number. Even when the R^2 is large, the sum of partial r^2 can be close to 0 if explanatory variables are highly correlated. On the contrary, this sum can be much larger than 1 if explanatory variables are independent and if the R^2 is large. Finally, the partial r^2 are highly dependent on which control variables are included in the regression, even if these are independent from the variable of interest. This property makes partial r^2 very hard to compare across regression models.[8] Given these three problems, I argue that partial r^2 should not

[8]The semi-partial r^2 solves this latter problem by comparing the variation explained exclu-

be used to measure the economic importance of variables in regressions.

4 New measures of economic importance

In this section, I define three interrelated measures of economic importance. First, the effect size of a variable x_i, denoted c_i, captures the magnitude of the effect of x_i on deviations in the dependent variable y, ceteris paribus. Second, the absolute importance of a variable x_i, denoted α_i, is a measure of the importance of x_i in the estimated model, which can be interpreted as a percentage. Finally, the relative importance of a variable x_i compared to another variable x_j is the ratio of their absolute importance, α_i/α_j.

I start the section by proposing a list properties or assumptions that the new measures of importance should ideally satisfy (Section 4.1). These are inspired by the discussion on the strengths and flaws of existing methods. The assumptions are then combined to form a measure of effect size, a measure of absolute economic importance, and a measure of relative economic importance (Section 4.2). Questions related to the operationalization and estimation of these measures are discussed in Sections 4.3 to 4.5.

4.1 Assumptions

The vector y is the weighted sum of $n+m$ variables $x_1, ..., x_n, x_{n+1}...., x_{n+m}$: $y = \beta_0 + \sum_{i=1}^{n+m} \beta_i x_i$. The elements of x_i are denoted x_{ik} $\forall k \in 1, ..., N$. The mean and standard deviation of x_i are denoted μ_i and σ_i respectively. The coefficient of correlation between x_i and x_j is denoted ρ_{ij}. The n variables $x_1, ..., x_n$ are observed, while the m variables $x_{n+1}...., x_{n+m}$ are unobserved and their weighted sum is denoted ϵ:

$$y = \beta_0 + \sum_{i=1}^{n+m} \beta_i x_i = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \epsilon. \tag{2}$$

sively by a variable to the total variation in the dependent variable.

I assume that ϵ is uncorrelated with observed variables, such that a simple OLS regression of y on $x_1...,x_n$ gives unbiased and consistent estimates of regression coefficients β_i.

Assumption 1 - relative importance and proportionality: *The relative importance of x_i compared to x_j is determined by the ratio of their effect sizes:*

$$\frac{\alpha_i}{\alpha_j} = \frac{c_i}{c_j} \; \forall i,j \in 1...n+m. \tag{3}$$

This assumption is consistent with the experimental evidence presented in section 2 and with the common practice of interpreting standardized beta coefficients.

Assumption 2 - percentage interpretation of absolute importance: *The sum of the measures of absolute importance is equal to 1.*

$$\left(\sum_{i=1}^{n+m} \alpha_i\right) = 1. \tag{4}$$

This assumption ensures that the measure of absolute economic importance α_i is easy to interpret. Thanks to this assumption, the measure α_i will be interpreted in percentage terms.

The first two assumptions establish the relationship between effect size, absolute importance, and relative importance. Equation (3) implies $\alpha_i = \alpha_j c_i/c_j \; \forall j \in 1,...,n+m$. Plugging these $n+m$ equations into equation (4) leads to:

$$\alpha_i = \frac{c_i}{\sum_{j=1}^{n} c_j} \forall i \in 1,...,n+m. \tag{5}$$

Equations (3) and (5) are expressing the absolute and relative importance of variables as a function of the effect sizes c_j. To simplify the discussion, the next assumptions are defined in terms of effect sizes. Following equations (3) and (5), they can be easily translated to refer to absolute and relative importance.

Assumption 3 - ceteris paribus: *The measure of effect size c_i associated with x_i relates to the partial effect of x_i on y_i. It only depends on β_i and on the x_{ik} $\forall k \in 1, ..., N$.*

In line with the interpretation of OLS regressions, the effect size of a variable x_i measures the ceteris paribus effect of x_i on y. This implies that c_i should not be affected by other variables and their coefficients, and by the coefficients of correlation ρ_{ij}: $\frac{\partial c_i}{\partial x_{jk}} = 0$, $\frac{\partial c_i}{\partial \beta_j} = 0$, $\frac{\partial c_i}{\partial \rho_{ij}} = 0$ $\forall j \neq i \in 1, ..., n+m$ and $\forall k \in 1, ..., N$.

This assumption is crucial because it limits the scope of the measure of importance α_i. It states that α_i focuses on partial derivatives and not on total derivatives. Total derivatives are often more interesting from theoretical and policy points of views. Empirically, however, calculating total derivatives is challenging in the absence of experimental or quasi-experimental data, because it requires identifying all causal relationships between explanatory variables. This is beyond the scope of most empirical works, especially if dynamics imply that it is always possible to find more remote causes, or if simultaneous relationships between explanatory variables lead to "chicken-and-egg" problems. While less ambitious, focusing on partial derivatives is more realistic and is in line with the current practice of comparing standardized beta coefficients.

Assumption 4 - invariance to linear transformations: *The measure of effect size c_i is invariant to linear transformations of explanatory variables.*

The effect size, and therefore the absolute economic importance of variables, should not be affected by linear changes in the unit of measurement of explanatory variables, for example from Celsius to Kelvin, or from kilometers to miles. Linear transformations in the unit of an explanatory variable x_i can entail the multiplication of all observations x_{ik} by a factor λ or the addition of a term Δ to all observations x_{ik} (or both). Let us consider these two possibilities in turn.

Any multiplication of x_i by a factor $\lambda \in \mathbb{R}$ will affect the statistical dispersion of x_i and affect β_i in inverse proportion, leaving $\beta_i x_i$ unchanged. According to this assumption, the effect size c_i can depend on measures of statistical dispersion

$d(\beta_j x_j)\ \forall\ j \in 1...n+m$, but not on the coefficients β_j or on measures of dispersion of the x_j separately.

In regression analysis, the addition of a term Δ to all observations x_{ik} of x_i will be compensated by a change in β_0 equivalent to $\beta_i\Delta$. Assumption 4 therefore implies that the measure of effect size c_i should not depend on the intercept β_0 and on measures of central tendencies of explanatory variables because these quantities are affected by linear transformations of the variables in the model. Thus, the measure of effect size c_i cannot depend on variables' mean, median or mode. The effect size c_i can however depend on measures of statistical dispersion that are location-invariant, i.e. such that $d(x_j + \Delta) = d(x_j)$. Examples of location-invariant measures of statistical dispersion include the standard deviation, the mean absolute deviation from a central point, the inter-quartile range, and the variance.

Assumption 5 - exclusion: *The measure of effect size c_i of variable x_i is equal to zero if β_i is equal to zero or if $x_{ik} = x_{il}\ \forall k, l \in 1, ..., N$.*

Assumption 6 - inclusion: *The measure of effect size c_i of variable x_i is greater than zero if β_i is different from zero and $\exists k, l \in 1, ..., N$ such that $x_{ik} \neq x_{il}$.*

Assumptions 5 and 6 are straightforward. Effect size is equal to zero if the regression coefficient is equal to zero or if the variable is a constant term; effect size is greater than zero otherwise. These assumptions imply that the formula defining the measure of effect size can only include terms that are a function of β or $d(x_i)$. As a corollary, the measure of effect size cannot include a constant term. Assumption 6 puts a further constraint on the function $d(x_i)$: it cannot depend uniquely on the quantiles of x_i.

Assumption 7 - linearity in regression coefficients: *The measure of effect size c_i of variable x_i is proportional to the size of the regression coefficient β_i:* $c_i \propto |\beta_i|$.

Consider the two DGPs $y_1 = \beta_0 + \beta_1 x_1 + ...$ and $y_2 = \beta_0 + \lambda\beta_1 x_1 + ...$ which are similar in all respects but the parameter λ. Assumption 7 quite logically states

that the effect size of x_1 in the second DGP is λ times the effect size of x_1 in the first DGP.

4.2 Effect size and economic importance

Taken together, Assumptions 3 to 7 imply that the effect size of variable x_i is measured as:

$$c_i = |\beta_i| d(x_i), \tag{6}$$

where $d(x_i)$ is a measure of statistical dispersion that is (1) location invariant, (2) equal to zero if $x_{ik} = x_{il}\ \forall k, l \in 1, ..., N$, and (3) greater than zero if $\exists k, l \in 1, ..., N$ such that $x_{ik} \neq x_{il}$. Following equation (5), we have that the absolute economic importance of variable x_i is measured as:

$$\alpha_i = \frac{c_i}{\sum_{j=1}^{n} c_j} = \frac{|\beta_i| d(x_i)}{\sum_{j=1}^{n} |\beta_j| d(x_j)}. \tag{7}$$

It is straightforward that this measure satisfies Assumption 2. From Assumption 1, we have that the relative economic importance of x_i compared to x_j is given by:

$$\frac{\alpha_i}{\alpha_j} = \frac{c_i}{c_j} = \frac{|\beta_i| d(x_i)}{|\beta_j| d(x_j)}. \tag{8}$$

Two decisions need to be taken before being able to estimate the measures of effect size, absolute importance, and relative importance derived in equations (6) to (8). The first decision relates to the choice of $d(x_i)$, the measure of statistical dispersion that is location invariant (Section 4.3). The second decision relates to how the unobservables - i.e. the error term ϵ - should be taken into account when estimating equation (7) (Section 4.4).

4.3 Measuring dispersion

Possible measures for $d(x_i)$ satisfying all assumptions include the standard deviation, denoted σ_i, and the mean absolute deviation from the mean, denoted δ_i.[9] The standard deviation is usually estimated by the corrected sample standard deviation:

$$s_i = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(x_{ik} - \bar{x_i})^2}.$$

This estimator is biased but consistent. The mean absolute deviation from the mean can be estimated as follows:

$$\hat{\delta}_i = \frac{1}{N}\sum_{k=1}^{N}|x_{ik} - \bar{x_i}|. \tag{9}$$

This estimator is also biased but consistent.[10]

Each of these measures has pros and cons. While standard deviations are widely used in the literature, they are difficult to visualize and interpret because their calculation involves a non-linear function of observations (it requires taking the square root of a sum of squares). In comparison, the mean absolute deviation from the mean is less popular but more intuitive, as it is a linear function of observations on each side of the mean. The mean absolute deviation simply measures the average distance of observations to the mean.

Let us consider the simple model $y = x_1 + x_2$ with $x_1 = (1, 0, 0, -1)$ and $x_2 = (2, 1, -1, -2)$. It is straightforward that $y = (3, 1, -1, -3)$. Quite logically, the means of the absolute deviations from the mean of x_1, x_2, and y are equal to 1/2, 3/2, and 2, respectively. Applying equation (7), we obtain that the absolute economic importance of x_1 and x_2 are equal to $\alpha_1 = 1/4$ or 25%, and $\alpha_2 = 3/4$

[9]The variance is excluded because of Assumption 7. The inter-quartile range and the median absolute deviation are excluded because of Assumption 6.

[10]The ratio $\hat{\delta}_i/\sqrt{1-1/n}$, which is an unbiased estimator of δ_i (Pham-Gia and Hung 2001), can also be used. This alternative estimator leads to the same measure of relative and absolute importance.

or 75%. These numbers are intuitively meaningful: they measure the proportion of absolute deviations in the dependent variable explained by each explanatory variable. In contrast, the standard deviations of x_1, x_2, and y are equal to $\sqrt{1/2}$, $\sqrt{5/2}$, and $\sqrt{5}$ respectively. The absolute economic importance of x_1 and x_2 is then given by: $\alpha_1 = 1/(1+\sqrt{5})$ or 31%, and $\alpha_2 = \sqrt{5}/(1+\sqrt{5})$ or 69%. These numbers are more difficult to interpret. This simple example shows that the mean absolute deviation from the mean is more appealing and meaningful than the standard deviation when it comes to assessing absolute economic importance.

Besides ease of interpretation, the mean absolute deviation from the mean has two other desirable properties. First, it is more robust to measurement errors because it gives less importance to extreme values. Second, although both s_i and $\hat{\delta}_i$ are asymptotically unbiased estimators of σ_i and of δ_i, respectively, $E(\hat{\delta}_i)$ converges toward δ_i faster than $E(s_i)$ does toward σ_i (Pham-Gia and Hung 2001).

For these reasons, the mean absolute deviation from the mean is adopted in what follows. It is however worth noting that the method works perfectly well with standard deviations. Differences between the two approaches will usually be relatively small, especially for continuous variables that are not too different from a normal or a uniform distribution. In fact, the two approaches give exactly the same results if all variables are normally distributed or if all variables are uniformly distributed.[11] Differences are expected to be more important for dummy variables whose means are close to 0 or close to 1, as standard deviations are much larger than mean absolute deviations for highly-skewed distributions.[12]

4.4 Accounting for the error term

How should unobservables be taken into account when assessing absolute economic importance? This question matters when estimating the denominator of

[11] This is because the ratio of the mean absolute deviation to the standard deviation is $\sqrt{2/\pi} = 0.798$ for normal distributions, and $\sqrt{3}/2 = 0.866$ for uniform distributions.

[12] For Bernoulli distributions (dummy variables), the ratio of the mean absolute deviation to the standard deviation is equal to $2\sqrt{p(1-p)}$ where p is the probability of observing a 1.

equation (7): how should one estimate the regression coefficients and dispersion of unobserved variables? I consider three possibilities, ranked in order of complexity.

Focusing on observables The first option is to focus only on explanatory variables included in the regression. In this case, the absolute economic importance of x_i can be easily estimated following equation (7). The β_j $\forall j \in 1, ..., n$ are estimated using regression analysis. The mean absolute deviation of each variable can be calculated following equation (9). The absolute importance of x_i is then estimated as:

$$\hat{\alpha}_i^1 = \frac{\hat{c_i}}{\sum_{j=1}^{n} \hat{c_j}} = \frac{|\hat{\beta}_i|\hat{\delta}_i}{\sum_{j=1}^{n} |\hat{\beta}_j|\hat{\delta}_j}. \tag{10}$$

The statistic $\hat{\alpha_i^1}$ associated with variable x_i is to be interpreted in percentage terms: it measures the proportion of ceteris paribus deviations in the predicted values of y explained by x_i.

Simple error term The previous approach may be deemed unsatisfactory because it completely ignores the error term, while researchers are often interested to estimate its importance. One other possibility is to assume that the error term consists of a single unobserved variable that can be estimated by the residuals of the regression. Following this assumption, the residuals are considered as one variable whose regression coefficient is equal to 1. The mean absolute deviation of residuals, denoted $\hat{\delta^\epsilon}$ can be estimated following equation (9).

According to this second approach, the absolute economic importance of a variable is given by:

$$\hat{\alpha}_i^2 = \frac{|\hat{\beta}_i|\hat{\delta}_i}{\hat{\delta^\epsilon} + \sum_{j=1}^{n} |\hat{\beta}_j|\hat{\delta}_j}. \tag{11}$$

Similarly, the absolute economic importance of unobservables is given by:

$$\hat{\alpha}_{\epsilon}^{2} = \frac{\hat{\delta^{\epsilon}}}{\hat{\delta^{\epsilon}} + \sum_{j=1}^{n} |\hat{\beta}_j| \hat{\delta}_j}. \tag{12}$$

The statistic $\hat{\alpha_i^2}$ measures, in percentage, the contribution of the variable x_i to deviations in the dependent variable, compared to the contributions of observables and of the error term.

In practice, the error term is often expected to be composed of a multitude of unobserved variables which are more or less important for predicting the dependent variable. The number and the variance of these unobserved variables is unknown. Assuming that the error term consists of a single unobserved variable therefore leads to an overestimation the contribution of observed variables when the error term is composed of multiple unobserved variables. In other words, the α_i^2 are upper-bound measures of the economic importance of explanatory variables.

Complex error term The composition of the error term can be approximated using information on observables. According to the third approach, unobserved variables are taken into account in the same way as variables explicitly included in the model. The contribution of each variable is therefore expressed as:

$$\alpha_i^3 = \frac{|\beta_i| \delta_i}{\sum_{j=1}^{n} |\beta_j| \delta_j + \sum_{k=1}^{m} |\beta_{n+k}| \delta_{n+k}}.$$

The challenge is to estimate the second term of the denominator, the sum $\sum_{k=1}^{m} |\beta_{n+k}| \delta_{n+k}$. Two pieces of information are unobserved and should be estimated: the number of variables composing the error term, m, and their effect size $c_{n+k} = |\beta_{n+k}| \delta_{n+k}$.

Different assumptions can be used to approximate these quantities. I assume that the unobserved variables are independent and normally distributed, implying that the following equalities hold for all $k \in n+1, ..., n+m$:[13]

[13]Note that for normal variables, the ratio of the mean absolute deviation to the standard deviation is $\sqrt{2/\pi}$.

$$\delta_k = \sqrt{\frac{2}{\pi}}\sigma_k$$
$$\delta^\epsilon = \sqrt{\frac{2}{\pi}}\sigma^\epsilon = \sqrt{\frac{2}{\pi}}\sqrt{\sum_{k=1}^{m}(\beta_{n+k}\sigma_{n+k})^2} = \sqrt{\sum_{k=1}^{m}(\beta_{n+k}\delta_{n+k})^2}.$$

I also assume that unobserved variables have the same dispersion: $\sigma_{n+k} = \sigma_{n+l}$ and $\delta_{n+k} = \delta_{n+l}\ \forall k, l \in 1, ..., m$. Finally, I assume that the effect size of each unobserved variable is equal to the average effect size of observables: $c_{n+k} = |\beta_{n+k}|\delta_{n+k} \approx \sum_{j=1}^{n} c_j/n = \bar{c}\ \forall k \in 1, ..., m$. Altogether, these assumptions imply:

$$\delta^\epsilon = \sqrt{m\beta_{n+k}\delta_{n+k}{}^2} = \sqrt{m\bar{c}^2} \Rightarrow m = (\frac{\delta^\epsilon}{\bar{c}})^2,$$

$$\sum_{k=1}^{m}\beta_{n+k}\delta_{n+k} = m\beta_{n+k}\delta_{n+k} = m\bar{c} = \frac{(\delta^\epsilon)^2}{\bar{c}}.$$

The adjusted measure of economic importance α_i^3, which accounts for the fact that the error term is expected to be composed of numerous unobserved variables, is therefore estimated as:

$$\hat{\alpha}_i^3 = \frac{|\hat{\beta}_i|\hat{\delta}_i}{\sum_{j=1}^{n}|\hat{\beta}_j|\hat{\delta}_j + \frac{(\hat{\delta^\epsilon})^2}{1/n\sum_{j=1}^{n}|\hat{\beta}_j|\hat{\delta}_j}}. \tag{13}$$

The importance of the error term is measured as:

$$\hat{\alpha}_\epsilon^3 = \frac{\frac{(\hat{\delta^\epsilon})^2}{1/n\sum_{j=1}^{n}|\hat{\beta}_j|\hat{\delta}_j}}{\sum_{j=1}^{n}|\hat{\beta}_j|\hat{\delta}_j + \frac{(\hat{\delta^\epsilon})^2}{1/n\sum_{j=1}^{n}|\hat{\beta}_j|\hat{\delta}_j}}. \tag{14}$$

When estimating $\bar{c}$ as $1/n\sum_{j=1}^{n}|\hat{\beta}_j|\hat{\delta}_j$, I recommend excluding irrelevant (insignif-

icant) variables, as these, if numerous, may artificially reduce the estimated value of $\bar{c}$, leading to an underestimation of the economic importance of observables.

How sensitive is $\hat{\alpha}_i^3$ to the supplementary assumptions we made on the composition of the error term? It depends. On the one hand, the assumption that the error term is composed of a sum of independent and normally distributed variables is not expected to affect $\hat{\alpha}_i^3$ much.[14] On the other hand, $\hat{\alpha}_i^3$ can dramatically change if we relax the assumption that the effect size of each unobserved variable is equal to the average effect size of observables. This assumption implicitly determines the number of variables composing the error term. At one extreme, the error term could be composed of a unique variable, in which case absolute economic importance should be measured by $\hat{\alpha}_i^2$. At the other extreme, the error term could be composed of an infinity of variables, in which case $\hat{\alpha}_i^3 \approx 0$ for each variable. The assumption that the effect size of each unobserved variable is equal to the average effect size of observables lies between these two extremes.

4.5 Remarks

The new method provides, in percentage terms, the proportion of deviations in the dependent variable explained by each explanatory variable, ceteris paribus. This new method is powerful, because it is simple to implement and easy to interpret. A few remarks are in order before presenting different applications of the method.

Which $\hat{\alpha_i^j}$ should be used? Importantly, the relative importance of variables does not depend on how the error term is accounted for. Therefore, the comparison of the relative importance of variables within a regression model is not affected by the choice of $\hat{\alpha_i^j}$.

If a research focuses on a single regression model, the measure $\hat{\alpha_i^1}$ should be preferred, as this simpler method does not require extra assumptions on the error term. The choice of $\hat{\alpha_i^j}$ matters when comparing different regression models. In

[14] For example, $\hat{\alpha}_i^3$ would not be much affected if the error term is composed of uniform variables, as the sum of n normal variables with a dispersion δ is only 8% larger than the sum of n uniform variables characterized by the same dispersion δ. This is especially true if the error term is small.

this case, I recommend using the measure $\hat{\alpha_i^2}$ with a full set of controls. The measure $\hat{\alpha_i^2}$ is convenient because it provides upper-bound estimates of the economic importance of variables included in the model while accounting for the error term. The measure $\hat{\alpha_i^2}$ should also be used in Monte-Carlo experiments, when the error term is generated as a unique variable. The addition of extra control variables uncorrelated with x_i but explaining y will always reduce the measures $\hat{\alpha_i^1}$ and $\hat{\alpha_i^2}$ associated with x_i. I therefore recommend applying these measures on a regression with a full set of control variables only. I recommend against $\hat{\alpha_i^3}$ because this latter measure relies on strong assumptions about the error term. The measure $\hat{\alpha_i^3}$ can underestimate or overestimate economic importance.[15]

IV or non-linear models The method presented in this paper only relies on regression coefficients and on the dispersion of explanatory variables. The method can therefore be used with IV estimates. It can also be easily adapted to non-linear functional forms. When calculating the magnitude c_i of the effect induced by a variable x_i and its nonlinear terms ${x_i}^2$,...,${x_i}^l$, all terms should be considered together: $c_i = d(\beta_{i1}x_i + \beta_{i2}{x_i}^2 + ... + \beta_{il}{x_i}^l)$. Assessing the importance of linear and nonlinear terms separately should be avoided because the contribution of each term would not be invariant to linear transformation of the variable. Similarly, the method can also be used to measure the importance of categorical variables. As for quadratic terms, all dummies related to a categorical variable should be considered together when calculating the importance of that variable.

Large sample properties The statistic $\hat{\alpha_i^j}$ is a function of regression coefficients and of the dispersion of the error term. It can hence be consistently estimated if regression coefficients are consistently estimated. However, this statistic is biased and inconsistent if regression coefficients are biased and inconsistent, for example because of omitted-variable bias, reverse causality, misspecification, or mismeasurement (see Section 5.1 and table A.10 for an illustration). The same is true for

[15]The addition of an extra control variable uncorrelated with x_i but explaining y may increase or decrease the measure $\hat{\alpha_i^3}$ associated with x_i, depending on whether the dispersion of the extra control is greater or smaller than the estimated dispersion of unobserved variables.

all measures of importance reviewed in this paper, including Shapley values.[16]

Small sample properties In small samples, the inclusion of variables that are irrelevant (not part of the true model) creates a downward bias in the estimated contributions of relevant variables. This bias is not only observed for the measure $\hat{\alpha_i^j}$, but also for the Shapley value and the partial-R^2. With small samples, there is a trade-off between the inclusion of potentially irrelevant controls and this downward bias. Significance thresholds can be used as criteria for excluding irrelevant variables and mitigate this source of bias. This problem disappears with large samples.

Multicollinearity In small samples, multicollinearity will bias all measures of importance (see Section 5.1 and table A.11 for an illustration). The bias is especially large for the partial R^2.[17] There are two easy solutions to mitigate this problem: (1) considering only one of the multicollinear variables, or (2) considering the multicollinear variables together when calculating their importance (if $\tilde{x_i}$ and $\breve{x_i}$ are the two multicollinear variable, their effect size is then calculated as $c_i = d(\beta_{i1}\tilde{x_i} + \beta_{i2}\breve{x_i})$). It is worth noting that this problem vanishes in large samples when it comes to the statistic $\hat{\alpha_i^j}$: the sum of the $\hat{\alpha_i^j}$ associated with multicollinear variables will converge to the actual importance of the variable in the population. This is not true for the Shapley value and the partial-R^2, which remain biased by multicollinearity even in large samples.

Sampling error and confidence intervals To provide a measure of sampling error, I recommend presenting bootstrap confidence intervals alongside estimates of $\hat{\alpha_i^j}$ (Wooldridge 2010). In the examples below, I use the Stata command *bsample* to randomly generate bootstrap samples. I then re-estimate statistics of interest using bootstrap samples and calculate the 0.025 and 0.975 quantiles to obtain 95% confidence intervals.

[16] Shapley values suffer from another, more serious, weakness: because Shapley values do not satisfy the "exclusion assumption", they are almost always biased and inconsistent, even in the absence of omitted variable bias, mismeasurement bias, reverse causality bias, and misspecification bias.

[17] The partial R^2 associated with the multicollinear variables will be extremely low.

Standard deviations Researchers who opt for standard deviations instead of mean absolute deviations can simply replace δ_j by σ_j $\forall j \in 1, ..., n$ in formulas (10) to (14). The impact of this change is expected to be marginal, except for highly-skewed variables (see Appendix D).

Variance and R^2 The method presented in this paper focuses on deviations in the dependent variable. Researchers willing to decompose the variance of the dependent variable instead of deviations can use the related measure $q_i^2 = Var(\beta_i x_i)/(Var(\epsilon) + \sum_{j=1}^{n} Var(\beta_j x_j))$ developed in Sterck (2018). Focusing on the variance instead of deviations has the advantage of simplifying calculations related to the importance of the error term.

5 Applications

5.1 Application 1: a simple data generating process

A simple monte-carlo experiment illustrates that the new method is more intuitive and easier to interpret than existing alternatives. I generated five normal variables x_1, x_2, x_3, x_4 and ϵ as $N(0, 1)$. These variables are combined to generate $y = x_1+2x_2+x_3+\epsilon$. Note that x_4 is absent from the DGP, implying that x_4 is irrelevant in regressions. The pairwise correlation between x_1, x_2, and x_4 is arbitrarily set at 0.5, while x_3 and ϵ are independent from the other variables. The number of observations was set at 10,000 to obtain precise estimates. In line with Section 2, the absolute economic importance of x_1, x_2, and x_3 is equal to 20%, 40%, and 20% respectively, while the error term, ϵ, captures 20% of deviations in y. The absolute economic importance of x_4 should be equal to 0 as this variable is irrelevant in the model.

Column 1 of table 1 presents the results of a regression of y on x_1, x_2, x_3, and x_4. As expected, coefficients associated with x_1, x_2, and x_3 are close to 1, 2, and 1 respectively. The coefficient of x_4 is close to 0 as this variable is irrelevant. Column 2 shows the results of a regression with standardized variables. The coefficient of x_1^*, 0.332, is not easy to interpret. It is approximately $1/\sqrt{9}$, where

9 is the variance of y. As shown in column 3, the mean value decomposition of Holgersson et al. (2014) leads to uninformative results when variables are centered in zero.[18] The Shapley values are shown in column 4. As explained previously, this method is flawed because it is highly sensitive to the correlation between variables included the model. The Shapley values of x_1 and x_3 are very different although these variables have the same regression coefficient and the same variance. While x_4 is irrelevant, its Shapley value is large - almost as large as the Shapley value of x_3. Partial and semi-partial r^2 are displayed in columns 5 and 6. These statistics cannot be intuitively interpreted because they are sensitive to the correlation between explanatory variables and because they do not add up to an easily interpretable number.

The outcome of the new method is presented in column 7. I focus on the measure α_i^2 because the error term is composed of a single variable. In line with the data generating process, the economic importance of x_1, x_2, x_3, and of the error term are estimated as: $\hat{\alpha_1^2} \approx 20.0\%$, $\hat{\alpha_2^2} \approx 39.9\%$, $\hat{\alpha_3^2} \approx 20.1\%$ and $\hat{\alpha_\epsilon^2} \approx 20.0\%$. The contribution of the irrelevant variable x_4 is approximately 0%. Bootstrap confidence intervals are narrow given the large number of observations in the generated dataset. In this example, the sum $|b_1^*| + |b_2^*| + |b_3^*| + |b_e^*|$ is equal to 1.66, implying that the standardized beta coefficients overestimate the percentage contributions of explanatory variables by 66%.

In online appendix, I study how results change when an extra regressor x_5 is added to the regression (table A.9). I show that the measure $\hat{\alpha_1^2}$ does not change if x_5 is irrelevant. The measure $\hat{\alpha_1^2}$ is slightly reduced if x_5 is relevant, and the reduction is maximal if x_5 explains 50% of the variance of the error term of the benchmark regression. In table A.10, I also show that all measures of importance reviewed in this paper are biased and inconsistent in the presence of omitted variable bias, mismeasurement bias, reverse causality bias, or misspecification bias. In table A.11, I show that multicollinearity is biasing all measures of importance in small samples. For $\hat{\alpha_i^2}$, this bias vanishes in large samples, in the sense that the sum of the $\hat{\alpha_1^2}$ associated with the two multicollinear variables is equal to 0.2,

[18] In the example, the means of y, x_1, x_2, and x_3 are equal to -0.005, -0.0002, 0.0002, 0.003.

which is the value of $\hat{\alpha_1^2}$ when there is no multicollinearity problem. Importantly, Shapley values and the partial-R^2 are affected by the multicollinearity problem even in large samples.

Table 1 – Linear regression model with non-orthogonal variables

	(1) β	(2) b^*	(3) $\beta_i \frac{\bar{x_i}}{\bar{y}}$	(4) Shapley values	(5) Partial r^2	(6) Semi-partial r^2	(7) New method α_i^2
x_1	0.997***	0.332***	0.045	0.226	0.398	0.073	0.200
	(0.004)	(0.001)					[0.198-0.201]
x_2	1.999***	0.664***	-0.076	0.460	0.726	0.294	0.399
	(0.004)	(0.001)					[0.397-0.401]
x_3	1.003***	0.334***	0.648	0.113	0.502	0.112	0.201
	(0.003)	(0.001)					[0.199-0.202]
x_4	0.001	0.000	0.001	0.090	0.000	0.000	0.000
	(0.004)	(0.001)					[0-0.002]
Residuals							0.200
							[0.199-0.201]
Constant	-0.002	-0.000	0.383				
	(0.003)	(0.001)					
N	100,000	100,000					
r2	0.89	0.89					

Standard errors in parentheses. Bootstrap confidence intervals (95%) in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Application 2: explaining emigration

What drives international migration to OECD countries? Are emigrants escaping dire prospects at home? Are they attracted by the perspective of higher incomes abroad? Are pre-existing diasporas shaping aspirations or facilitating movement? In a globalized world, are geographic and cultural distances playing an important role? How do these factors vary with the education level of the migrants?

In order to answer these important questions, I use the dyadic data of Dao et al. (2018) and extend their analysis. In line with Dao et al. (2018) and Docquier et al. (2014), I consider three categories of dependent variables:

1. *Migration rates:* This variable captures net migration flows of working-age population to OECD countries over the 2000-2010 period. For each country

of origin i and each OECD destination country j, net migration flows are proxied by the difference between the bilateral migrant stocks in 2010 and 2000 (denoted M_{ij}), expressed in percentage of the resident population at origin in 2000 (denoted N_i): $m_{ij} = M_{ij}/N_i$.[19]

2. *Potential migration rates:* This dependent variable is the sum of net migration flows M_{ij} and of the stock of aspiring migrants (denoted A_{ij}), expressed in percentage of the resident population at origin in 2000: $p_{ij} = (M_{ij} + A_{ij})/N_i$.[20]

3. *Migration realization rates:* This dependent variable captures the proportion of potential migrants ($M_{ij} + A_{ij}$) realizing their aspiration: $r_{ij} = M_{ij}/(M_{ij} + A_{ij}) = m_{ij}/p_{ij}$.

These three dependent variables are log-transformed to improve the fit of the regressions (migration data is highly skewed) and to capture the interesting relationship $\log(m_{ij}) = \log(p_{ij}) + \log(r_{ij})$. For each dependent variable, I distinguish (aspiring) emigrants with college education from those with lower levels of education.

Explanatory variables are grouped in four categories:

1. *Pre-existing diaspora:* the size of the migrant network of origin country i in destination country j is measured by the log of the total stock of migrants from i in j in the year 2000 divided by the population size of i in 2000. As in Dao et al. (2018), this variable is instrumented with its 10-year lag to minimize endogeneity.[21]

2. *Gravity variables:* Geographical distance between origin and destination countries is captured by the log of their distance and by a dummy equal

[19] As in Dao et al. (2018), only dyads with positive potential migration flows and realization rates strictly between 0 and 1 are considered in the analysis.

[20] The stock of aspiring migrants A_{ij} is calculated using the following two questions included in the 2007–2010 Gallup World Poll surveys: (i) *"Ideally, if you had the opportunity, would you like to move to another country, or would you prefer to continue living in this country?"* (ii) *"To which country would you like to move?"*

[21] In line with Dao et al. (2018)'s choice to exclude dyads with no migration flows between 2000 and 2010, dyads with no migration stock in 1990 or 2000 are excluded from the analysis to minimize the influence of inconsistent observations.

to 1 if they are contiguous. Cultural distance is proxied by their genetic distance and by dummy variables equal to 1 if they share a common language or a colonial link. I also control for changes in the restrictiveness of migration policies of the destination country towards the origin country between 2000 and 2010.

3. *Socio-economic conditions in the origin country:* Push factors are captured by the GDP per capita of the origin country in 2000 (log) and its square, by a measure of education quality, and by a measure of openness to trade. GDP per capita in the origin country also accounts for the financial capacity of potential migrants. I control for the size of the population (log), and for the share of the population in the origin country aged between 15 and 24 in 2000. This latter variable is a proxy for the population in the age of migration.

4. *Characteristics of the destination country:* I use the GDP per capita of the destination country in 2000 (log) to proxy for its attractiveness, and the population size of the destination country in 2000 (log) to proxy for its hosting capacity.

Results of IV regrsesions are shown are shown in Table 2 for all migrants, in Table 3 for less educated migrants, and in Table 4 for college graduate migrants. In each table, I distinguish migration rates (columns 1 and 2) from potential migration rates (columns 3 and 4), and migration realization rates (columns 5 and 6). Regression coefficients are shown in columns 1, 3, and 5, while the metrics $\hat{\alpha}_i^2$ are shown in columns 2, 4, and 6. I focus here on $\hat{\alpha}_i^2$ because I aim at comparing the importance of variables across regression models. Similar conclusions are reached with the measures $\hat{\alpha}_i^1$ and $\hat{\alpha}_i^3$, as these are proportional to $\hat{\alpha}_i^2$ (see online Appendix).

The size of the diaspora in destination countries stands out as the most important variable across all models. The size of the network has a strong positive effect on the aspiration and realization rates of both college-educated and less-educated individuals. This variable alone explains as much as 36% of deviations in overall migration rates.

By comparison, gravity variables appear to be relatively unimportant. Geographic distance is responsible for 5.4% of deviations in migration rates. This effect is mainly driven by less educated migrants, who are less likely to aspire to migrate to distant countries, and less likely to realize these aspirations. Sharing a common language is a relatively important factor for college graduates, but not for less educated individuals. The impact of genetic distance is ambiguous. On the one hand, genetic distance has a positive effect on migration aspirations, especially for college graduates, suggesting that educated individuals are looking for exoticism. On the other hand, genetic distance appears to be an important barrier for less educated individuals who aspire to migrate. The resulting impact of genetic distance on overall migration rates is negative, but relatively unimportant. Sharing a border, having colonial links, and the restrictiveness of migration policies at destination appear to be relatively unimportant factors.

Among the socio-economic characteristics of origin countries, only GDP per capita and population size appear to be important drivers of migration. While GDP per capita and its square are not significantly associated with aspiration rates (p-value of joint F-test= 0.73), these variables jointly explain about 16% of deviations in migration realization rates. The shape of the relationship between the two variables is an inverted U-shape that peaks at \$3,226 per capita in PPP value. The positive section of the relationship is mostly driven by less educated individuals, suggesting that their limited financial capacity impedes the realization of their migration aspirations. The strong inverted U-shaped relationship between GDP per capita and realization rates implies that overall migration rates also exhibit an inverted U-shaped relationship with GDP per capita, which is maximal at \$2,762 per capita in PPP value. GDP per capita explains 11.7% of deviations in overall migration rates. The coefficient associated with population size of origin countries is negative and statistically significant, suggesting that residents of larger countries are less likely to migrate abroad. This effect explains about 6% of deviation in overall migration rates. The degree of openness to trade and the quality of education of the origin countries appear to be relatively unimportant.

While people aspire to migrate to richer countries, this effect relatively unim-

portant compared to the effect of other variables. Also, realization rates are not significantly correlated with the GDP per capita of destination countries. The resulting effect of destination countries' GDP per capita on overall migration rates is therefore positive, but the effect is rather small, especially compared to the effect the GDP per capita in origin countries.

In conclusion, migration rates mainly depend on two variables: the size of the diaspora in destination countries (36.1%), and GDP per capita in origin countries (11.7%). This latter relationship is non linear: migration rates are increasing in GDP per capita below $2,762 per capita in PPP value, and decreasing in GDP per capita past this threshold. The increasing section of the inverted-U is mainly driven the very low realization rates of less educated individuals living in poor countries. Importantly, these results focus on the partial effects of explanatory variables on migration flows between 2000 and 2010. Therefore, these results are uninformative about total derivatives and about dynamics before 2000.[22]

5.3 Application 3: explaining fertility

The key determinants of fertility are well-known (e.g. age, marital status, education, child mortality, population density). The literature is however silent on their relative economic importance. What are the most important determinants of fertility? Is it age, education, martial status, child mortality, or population density? This important question can be answered by applying the new method to the recent analysis of de la Croix and Gobbi (2017).

The research de la Croix and Gobbi (2017) examines whether population density explains within-country variations in the number of birth per women. The authors emphasize that understanding the determinants of fertility is key to improve long-run population projections, and better assess the sustainability of contemporary societies. In their preferred specification, de la Croix and Gobbi (2017) propose to minimize omitted-variable bias and attenuation bias by instrumenting

[22]Studying total derivatives and dynamics before 2000 would require identifying (or hypothesizing) all causal relationships between explanatory variables as well as determining how these causal relationships evolved over time. This is beyond the scope of this paper.

Table 2 – Determinants of migration by dyad

	All migrants					
	Migration rate		Aspiration rate		Realization rate	
	β_i	α_i^2	β_i	α_i^2	β_i	α_i^2
	(1)	(2)	(3)	(4)	(5)	(6)
Network (% pop., in log)	0.650***	36.056	0.418***	28.273	0.232***	17.707
	(0.027)	[31.7-38.5]	(0.022)	[24-30.7]	(0.029)	[13.8-20.9]
Geo. Dist. (log)	-0.198***	5.394	-0.121***	4.001	-0.078*	2.908
	(0.044)	[3.8-6.7]	(0.046)	[2-6]	(0.042)	[0.7-4.7]
Contiguity	-0.295**	0.710	-0.438***	1.287	0.144	0.477
	(0.150)	[0-1.4]	(0.116)	[0.5-2.2]	(0.138)	[0-1.6]
Genetic Dist.	-0.166**	3.198	0.134	3.151	-0.300***	7.971
	(0.068)	[1.5-4.6]	(0.088)	[0.7-5.6]	(0.092)	[5.3-9.9]
Com. Lang.	0.572***	4.785	0.537***	5.477	0.035	0.399
	(0.089)	[3.9-5.9]	(0.101)	[3.3-7.1]	(0.106)	[0.1-2.2]
Colonial Link	0.087	0.168	-0.107	0.252	0.195	0.516
	(0.116)	[0-0.5]	(0.139)	[0-1]	(0.145)	[0-1.2]
Pol. restr.	-0.057	0.277	-0.227**	1.345	0.170*	1.139
	(0.095)	[0-1.1]	(0.092)	[0.4-2.3]	(0.089)	[0.2-2.2]
GDP/cap	2.903***	11.713	-0.459	1.285	3.361***	15.992
	(0.639)	[8.9-13.6]	(0.606)	[0.5-4.9]	(0.761)	[12.8-18.7]
GDP/cap Sq.	-0.183***		0.025		-0.208***	
	(0.038)		(0.034)		(0.045)	
Educ. Quality	-0.002	0.418	-0.009	2.756	0.007	2.541
	(0.007)	[0-2.4]	(0.010)	[0.2-5.6]	(0.010)	[0.1-7]
Import Tariff	-0.004	0.580	-0.002	0.453	-0.001	0.287
	(0.008)	[0-2.6]	(0.010)	[0-3.1]	(0.011)	[0.1-3.7]
Population (log)	-0.168***	5.997	-0.069**	3.007	-0.099**	4.863
	(0.035)	[4.4-7.3]	(0.031)	[1.2-4.8]	(0.044)	[1.9-7]
Pop 15-24 (% pop.)	-0.000	0.038	0.005	0.602	-0.006	0.734
	(0.021)	[0.1-3.2]	(0.025)	[0.1-4.3]	(0.028)	[0.1-5]
GDP/cap	0.421***	2.577	0.488***	3.635	-0.067	0.560
	(0.119)	[1.2-3.6]	(0.105)	[2-4.8]	(0.111)	[0-2.1]
Population (log)	0.175***	6.132	0.331***	14.115	-0.156***	7.516
	(0.026)	[4.2-7.8]	(0.034)	[12.1-15.8]	(0.037)	[5-9.6]
Residuals		21.956		30.363		36.390
		[19.5-23.1]		[27.7-31.5]		[31.8-39.5]
R-squared	0.77		0.62		0.29	
N. of obs	1358		1358		1358	
F-stat 1^{st} stage	1974.58		1974.58		1974.58	

Robust standard errors, clustered by country of origin, in parentheses. Bootstrap confidence intervals (95%) in square brackets. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean of the variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3 – Determinants of migration by dyad

	Less educated migrants					
	Migration rate		Aspiration rate		Realization rate	
	β_i	α_i^2	β_i	α_i^2	β_i	α_i^2
	(1)	(2)	(3)	(4)	(5)	(6)
Network (% pop., in log)	0.693***	34.869	0.422***	27.213	0.271***	16.807
	(0.032)	[31.2-38.1]	(0.021)	[23.4-30]	(0.032)	[13.2-19]
Geo. Dist. (log)	-0.339***	8.523	-0.167***	5.366	-0.173***	5.345
	(0.057)	[6.6-10.7]	(0.050)	[3.5-7.1]	(0.049)	[2-7.6]
Contiguity	-0.399	0.862	-0.428***	1.185	0.029	0.077
	(0.246)	[0-2]	(0.124)	[0.4-2.1]	(0.248)	[0-1.6]
Genetic Dist.	-0.122	2.231	0.103	2.408	-0.226**	5.068
	(0.088)	[0-4.1]	(0.092)	[0.4-4.7]	(0.104)	[1.5-7.7]
Com. Lang.	0.276**	2.212	0.523***	5.371	-0.247*	2.440
	(0.122)	[0.5-3.9]	(0.115)	[3.2-7.1]	(0.137)	[0.2-4.4]
Colonial Link	-0.230	0.455	-0.255	0.646	0.025	0.061
	(0.193)	[0-1.1]	(0.170)	[0-1.5]	(0.138)	[0-0.9]
Pol. restr.	-0.057	0.258	-0.197*	1.150	0.140	0.788
	(0.115)	[0-1.5]	(0.107)	[0.1-2]	(0.115)	[0.1-2.1]
GDP/cap	3.096***	12.320	-0.995	2.479	4.091***	15.426
	(0.784)	[9.2-15.3]	(0.633)	[0.7-4.8]	(0.920)	[12.2-18.1]
GDP/cap Sq.	-0.198***		0.055		-0.254***	
	(0.047)		(0.035)		(0.055)	
Educ. Quality	-0.011	2.567	-0.017*	5.111	0.006	1.752
	(0.008)	[0.3-4.4]	(0.010)	[1.2-8.2]	(0.010)	[0.1-6.3]
Import Tariff	-0.010	1.579	-0.000	0.089	-0.010	1.861
	(0.009)	[0.1-3.7]	(0.011)	[0.1-3.2]	(0.012)	[0.1-4.1]
Population (log)	-0.159***	5.222	-0.064*	2.692	-0.095*	3.849
	(0.045)	[3.5-7]	(0.035)	[0.6-4.2]	(0.053)	[1.7-5.9]
Pop 15-24 (% pop.)	-0.008	0.719	-0.008	0.934	0.000	0.012
	(0.027)	[0.1-4]	(0.027)	[0.1-4.3]	(0.033)	[0.1-4.1]
GDP/cap	0.077	0.435	0.263**	1.894	-0.185	1.286
	(0.167)	[0.1-1.6]	(0.106)	[0.6-3.6]	(0.147)	[0-2.9]
Population (log)	0.071***	2.319	0.317***	13.248	-0.246***	9.888
	(0.027)	[0.8-4]	(0.035)	[10.6-15.3]	(0.040)	[7.4-11.7]
Residuals		25.428		30.213		35.340
		[23-27.1]		[27.1-31.8]		[30.8-37.3]
R-squared	0.68		0.60		0.28	
N. of obs	1198		1198		1198	
F-stat 1^{st} stage	1783.23		1783.23		1783.23	

Robust standard errors, clustered by country of origin, in parentheses. Bootstrap confidence intervals (95%) in square brackets. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean of the variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 – Determinants of migration by dyad

	Migration rate		College graduates Aspiration rate		Realization rate	
	β_i	α_i^2	β_i	α_i^2	β_i	α_i^2
	(1)	(2)	(3)	(4)	(5)	(6)
Network (% pop., in log)	0.648***	29.236	0.314***	21.335	0.334***	23.913
	(0.034)	[25.9-32.4]	(0.026)	[18.2-24.5]	(0.029)	[19.6-27.9]
Geo. Dist. (log)	-0.142***	3.279	-0.091**	3.170	-0.051	1.863
	(0.049)	[1.4-4.9]	(0.039)	[1.1-5.1]	(0.043)	[0.3-4.3]
Contiguity	-0.653***	1.795	-0.475***	1.968	-0.178	0.776
	(0.121)	[1-2.4]	(0.137)	[0.6-2.8]	(0.124)	[0-1.8]
Genetic Dist.	0.188*	2.787	0.298***	6.644	-0.110	2.578
	(0.111)	[1.4-4.6]	(0.084)	[4.1-8.9]	(0.100)	[0.5-5.5]
Com. Lang.	0.847***	6.140	0.467***	5.096	0.380***	4.374
	(0.128)	[4.4-7.8]	(0.104)	[3.1-7.2]	(0.101)	[2.3-6.2]
Colonial Link	0.250	0.438	0.156	0.412	0.094	0.261
	(0.159)	[0-1.2]	(0.126)	[0-1.2]	(0.135)	[0-1.2]
Pol. restr.	0.118	0.569	-0.043	0.317	0.161**	1.237
	(0.088)	[0-1.4]	(0.073)	[0-1.9]	(0.070)	[0-2.1]
GDP/cap	-0.304	12.629	-1.433*	4.145	1.129	17.240
	(0.799)	[9-15.7]	(0.824)	[2.1-7.5]	(0.836)	[13.5-20.2]
GDP/cap Sq.	-0.011		0.076		-0.086*	
	(0.046)		(0.047)		(0.049)	
Educ. Quality	-0.013	2.541	-0.016	4.737	0.003	0.958
	(0.011)	[0.1-5]	(0.011)	[2.2-8.9]	(0.011)	[0.1-4.7]
Import Tariff	0.025*	3.298	0.015	3.036	0.010	2.034
	(0.013)	[1.6-5.4]	(0.011)	[0.6-5.7]	(0.012)	[0.1-4.7]
Population (log)	-0.213***	6.555	-0.084**	3.875	-0.129***	6.318
	(0.043)	[4.8-8.4]	(0.042)	[2-6.3]	(0.044)	[4.3-8.1]
Pop 15-24 (% pop.)	-0.015	1.193	0.046	5.407	-0.062**	7.589
	(0.032)	[0.2-3.1]	(0.033)	[0.2-9.1]	(0.030)	[4.5-11.3]
GDP/cap	1.005***	4.680	0.814***	5.709	0.191	1.412
	(0.124)	[3.2-5.7]	(0.106)	[4.2-7.3]	(0.131)	[0.1-2.7]
Population (log)	0.190***	5.148	0.192***	7.836	-0.002	0.086
	(0.036)	[3.5-7.1]	(0.033)	[6-10.2]	(0.033)	[0-2.3]
Residuals		19.712		26.314		29.363
		[18-21]		[24.4-27.8]		[25.8-32.6]
R-squared	0.79		0.63		0.48	
N. of obs	965		965		965	
F-stat 1^{st} stage	1385.26		1385.26		1385.26	

Robust standard errors, clustered by country of origin, in parentheses. Bootstrap confidence intervals (95%) in square brackets. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean of the variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

population density with the distance to buildings and cities belonging to UNESCO World Heritage Sites constructed between the Neolithic Revolution and 1900. Their analysis uncovers a statistically significant relationship between population density and fertility. Their extremely rich dataset can be used to push the analysis further, and draw a number of inferences regarding the economic importance of different channels.

Table 5 replicates the analysis of de la Croix and Gobbi (2017). It shows the results of an OLS regression (column 1) and an IV regression (column 3) with individual level data. Because I analyze a single regression model, I focus on the metrics $\hat{\alpha}_i^1$, which are shown in columns 2 and 4.

Because of the large number of observations in the dataset, most regression coefficients are statistically significant. Looking at the " stars" does not tell us much about the importance of the different channels. The new method allows comparing the economic importance of variables included in the regression. Perhaps not surprisingly, age is by far the most important determinant of the number of birth per woman. Age dummies explain about 46.6% of deviations in the dependent variable due to observables. Being married is identified as the second most important factor, explaining about 11% of deviations. Women's education is also a key determinant of fertility, responsible for about 10.2% of explained deviations in the dependent variable. The education level of women living in the same cluster also seems to matter when looking at the results of the OLS regression. The estimated importance of this factor diminishes substantially when population density is instrumented. Results show that women compensate the death of a child by increasing their fertility rate. This factor captures about 7.9% of the deviations explained by variables included in the model.

Population density seems to be a relatively unimportant determinant of fertility when looking at the results of the OLS regression. The regression coefficient of population density and its estimated importance triple when the variable is instrumented. In the IV specification, population density explains 10% of deviations

predicted by regressors.[23]

Despite being statistically significant, several variables appear to be unimportant in explaining birth per women when other factors are controlled for. The measure of GDP per capita (proxied by night-time light satellite data), child mortality the cluster level, land quality, and the distance to large bodies of water only explain a tiny fraction of deviations in fertility (1.6%, 1.1%, 0.9%, and 1.1% respectively).

The new method enriches the insightful analysis of de la Croix and Gobbi (2017), by assessing the economic importance of different channels of fertility, ceteris paribus. This example illustrates the importance of assessing economic importance when analyzing large databases, with which even minor differences and unimportant factors are likely to be statistically significant. It is worth noting that the results above focus on the partial effects of explanatory variables. An in-depth theoretical and empirical analysis of the causal relationships between explanatory variables would be needed to study total derivatives as well as dynamics over time. This is beyond the scope of this paper.

6 Conclusion

In March 2016, the American Statistical Association released an important report on "Statistical Significance and P-Values." While this report rightly highlights that *"a p-value, or statistical significance, does not measure the size of an effect or the importance of a result."*, it does not explain what importance means, nor how it should be measured. In this paper, I have explored these two questions. I have argued that an empirical result is economically important if two conditions are met: (1) the dependent variable under scrutiny must be economically relevant, and (2) the effect of the explanatory variable on the dependent variable must be sizable. This paper focused on this latter, more objective, aspect of economic importance, assuming that the dependent variable is economically relevant. Importantly, the

[23]It is beyond the scope of this paper to explain the large difference between the OLS and the IV coefficients.

Table 5 – Determinants of women's total number of births

	OLS		IV	
	β	α_i^1	β	α_i^1
	(1)	(2)	(3)	(4)
ln(1+density)	-0.045***	2.552	-0.158***	8.849
	(0.002)	[2.4-2.7]	(0.015)	[8-9.6]
Married	0.924***	11.032	0.921***	10.770
	(0.009)	[10.9-11.2]	(0.009)	[10.6-11]
Mean marriage	-0.174***	0.870	-0.375***	1.840
	(0.032)	[0.7-1.1]	(0.043)	[1.6-2.1]
Child mortality	1.961***	7.867	1.960***	7.704
	(0.024)	[7.7-8]	(0.024)	[7.5-7.9]
Mean child mortality	0.608***	1.068	0.602***	1.036
	(0.086)	[0.8-1.2]	(0.090)	[0.8-1.3]
ln(GDP per capita)	-0.047***	1.616	-0.006	0.194
	(0.005)	[1.4-1.8]	(0.008)	[0-0.5]
Woman's education	-0.019***	10.168	-0.020***	9.983
	(0.002)	[10-10.4]	(0.002)	[9.7-10.3]
Woman's education squared	-0.005***		-0.005***	
	(0.000)		(0.000)	
Education in cluster	-0.041***	6.554	0.001	2.688
	(0.005)	[6.3-6.8]	(0.008)	[2.1-3.3]
Education in cluster squared	-0.002***		-0.002***	
	(0.000)		(0.000)	
Caloric suitability index	0.009***	0.895	0.011***	1.033
	(0.001)	[0.7-1.1]	(0.002)	[0.8-1.2]
Distance to water	0.025***	1.130	-0.017**	0.730
	(0.003)	[1-1.3]	(0.007)	[0.5-1]
Age dummies	Yes	46.558	Yes	45.615
		[46.2-46.9]		[45.2-46.1]
Country dummies	Yes	9.690	Yes	9.559
		[9.6-9.8]		[9.3-9.7]
Observations	490,669		490,669	
R^2	0.605		0.601	

Robust standard errors, clustered at the cluster level, in parentheses. Bootstrap confidence intervals (95%) in square brackets. All specifications include age dummies and country fixed effects. OLS regressions are shown in columns 1 to 2. IV regressions are presented in columns 3 and 4. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean of the variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

objective assessment of whether an effect is large should not be confounded with the subjective assessment of whether an effect is surprising or interesting.

I have shown that the tools commonly used to assess economic importance are imperfect. Building on the strengths and weaknesses of existing methods, I constructed different measures of absolute and relative economic importance of variables, as well as residuals in linear regressions. The measures focus on certeris paribus effects (i.e. partial derivatives). They are easy to build and interpret. They are objective, as they do not depend on a researcher's value judgment nor on the label of explanatory variables. The new measures are particularly relevant when the phenomenon studied is caused by a multitude of variables that are not manipulated by the researcher. By contrast, cost-benefit analysis is more appropriate for empirical studies focusing on the effect of one or few interventions in a controlled setting (e.g. RCTs and other impact evaluation methods). The new method was applied to study the causes of international migration to OECD countries, and explore determinants of fertility.

To be sure, the new approach introduced in this paper should not be regarded as a miracle solution that can be applied without careful thinking. First, because the statistics $\hat{\alpha_i^j}$ are inconsistent in the presence of omitted variable bias, mismeasurement bias, reverse causality bias, and misspecification bias. Second, because the statistics $\hat{\alpha_i^j}$ focus on partial derivatives, and are therefore uninformative about total derivatives. Theory is therefore particularly important because it helps identifying and solving possible endogeneity issues, but also because it allows going beyond partial derivatives to explore total derivatives and dynamics.

This paper and its applications highlight the importance of having a credible and intuitive measure of economic importance for descriptive studies that aim at understanding the causes of a phenomenon, but that do not aim to directly affect this phenomenon. As research analyzing large datasets is more and more frequent, evaluating economic importance is indeed becoming as - if not more - important than assessing statistical significance. This research proposed a simple and intuitive method to measure economic importance that can usefully complement

standard measures of statistical significance.

References

Achen, C. H. (1982). *Interpreting and using regression*, Volume 29. London, UK: Sage.

Aghion, P., P. Bacchetta, R. Ranciere, and K. Rogoff (2009). Exchange rate volatility and productivity growth: The role of financial development. *Journal of monetary economics 56*(4), 494–513.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2015). Ethnic inequality. *Journal of Political Economy 123*(3), 547–724.

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*(1), 151–184.

Arbath, C. E., Q. H. Ashraf, O. Galor, and M. Klemp (2018). Diversity and conflict. Technical report, Brown University, Department of Economics.

Ashraf, Q. and O. Galor (2013). The "out of Africa" hypothesis, human genetic diversity, and comparative economic development. *The American Economic Review 103*(1), 1–46.

Bring, J. (1994). How to standardize regression coefficients. *The American Statistician 48*(3), 209–213.

Brodeur, A., M. Lé, M. Sangnier, and Y. Zylberberg (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics 8*(1), 1–32.

Budescu, D. V. (1993). Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression. *Psychological bulletin 114*(3), 542.

Camerer, C. F., A. Dreber, E. Forsell, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, J. Almenberg, A. Altmejd, T. Chan, et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science 351*(6280), 1433–1436.

Chevan, A. and M. Sutherland (1991). Hierarchical partitioning. *The American Statistician 45*(2), 90–96.

Dao, T. H., F. Docquier, C. Parsons, and G. Peri (2018). Migration and development: Dissecting the anatomy of the mobility transition. *Journal of Development Economics 132*, 88–101.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological bulletin 69*(3), 161.

de la Croix, D. and P. E. Gobbi (2017). Population density, fertility, and demographic convergence in developing countries. *Journal of Development Economics 127*(Supplement C), 13 – 24.

Djankov, S., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2008). The law and economics of self-dealing. *Journal of financial economics 88*(3), 430–465.

Docquier, F., G. Peri, and I. Ruyssen (2014). The cross-country determinants of potential and actual migration. *International Migration Review 48*, S37–S99.

Docquier, F. and H. Rapoport (2012). Globalization, brain drain, and development. *Journal of Economic Literature 50*(3), 681–730.

Dustmann, C. and A. Okatenko (2014). Out-migration, wealth constraints, and the quality of local amenities. *Journal of Development Economics 110*, 52–63.

Franco, A., N. Malhotra, and G. Simonovits (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science 345*(6203), 1502–1505.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine 27*(15), 2865–2873.

Grömping, U. (2015). Variable importance in regression models. *Wiley Interdisciplinary Reviews: Computational Statistics 7*(2), 137–152.

Grömping, U. et al. (2006). Relative importance for linear regression in r: the package relaimpo. *Journal of statistical software 17*(1), 1–27.

Hart, A. G. (1965). Capital appropriations and the accelerator. *The Review of Economics and Statistics*, 123–136.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2018). The global distribution of economic activity: nature, history, and the role of trade. *The Quarterly Journal of Economics 133*(1), 357–406.

Holgersson, H., T. Norman, and S. Tavassoli (2014). In the quest for economic significance: assessing variable importance through mean value decomposition. *Applied Economics Letters 21*(8), 545–549.

Johnson, J. W. and J. M. LeBreton (2004). History and use of relative importance indices in organizational research. *Organizational research methods 7*(3), 238–257.

Kruskal, W. (1987). Relative importance by averaging over orderings. *The American Statistician 41*(1), 6–10.

Kruskal, W. and R. Majors (1989). Concepts of relative importance in recent scientific literature. *The American Statistician 43*(1), 2–6.

Lindeman, R. H. (1980). Introduction to bivariate and multivariate analysis. Technical report.

Lipovetsky, S. and M. Conklin (2001). Analysis of regression in game theory approach. *Applied Stochastic Models in Business and Industry 17*(4), 319–330.

Lipsey, R. G. and M. Steuer (1961). The relation between profits and wage rates. *Economica 28*(110), 137–155.

Luskin, R. C. (1991). Abusus non tollit usum: standardized coefficients, correlations, and r 2s. *American Journal of Political Science*, 1032–1046.

Malmendier, U. and S. Nagel (2011). Depression babies: do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics 126*(1), 373–416.

Manchin, M. and S. Orazbayev (2018). Social networks and the intention to migrate. *World Development 109*, 360–374.

Maniadis, Z., F. Tufano, and J. A. List (2017). To replicate or not to replicate? exploring reproducibility in economics through the lens of a model and a pilot study. *The Economic Journal 127*(605).

McCloskey, D. N. and S. T. Ziliak (1996). The standard error of regressions. *Journal of Economic Literature 34*(1), 97–114.

Michalopoulos, S. (2012). The origins of ethnolinguistic diversity. *The American Economic Review 102*(4), 1508–1539.

Miguel, E., C. Camerer, K. Casey, J. Cohen, K. M. Esterling, A. Gerber, R. Glennerster, D. P. Green, M. Humphreys, G. Imbens, et al. (2014). Promoting transparency in social science research. *Science 343*(6166), 30–31.

Nakagawa, S. and I. C. Cuthill (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological reviews 82*(4), 591–605.

Nunn, N. and N. Qian (2014). Us food aid and civil conflict. *American Economic Review 104*(6), 1630–66.

Olken, B. A. (2015). Promises and perils of pre-analysis plans. *The Journal of Economic Perspectives 29*(3), 61–80.

Oster, E. (2014). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*.

Pham-Gia, T. and T. Hung (2001). The mean and median absolute deviations. *Mathematical and Computer Modelling 34*(7-8), 921–936.

Pritchett, L. and J. Sandefur (2015). Learning from experiments when context matters. *The American Economic Review 105*(5), 471.

Satyanath, S., N. Voigtländer, and H.-J. Voth (2017). Bowling for fascism: Social capital and the rise of the nazi party. *Journal of Political Economy 125*(2), 478–526.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games 2*(28), 307–317.

Shea, J. (1997). Instrument relevance in multivariate linear models: A simple measure. *Review of Economics and Statistics 79*(2), 348–352.

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *Journal of Economic Inequality*, 1–28.

Spolaore, E. and R. Wacziarg (2009). The diffusion of development. *The Quarterly Journal of Economics 124*(2), 469–529.

Sterck, O. (2018). On the economic importance of the determinants of long-term growth. *CSAE working paper 2018-20*.

Stufken, J. (1992). On hierarchical partitioning. *American Statistician 46*(1), 70–71.

Valencia Caicedo, F. (2018). The mission: Human capital transmission, economic persistence, and culture in south america. *The Quarterly Journal of Economics 134*(1), 507–556.

Vivalt, E. (2015). Heterogeneous treatment effects in impact evaluation. *The American Economic Review 105*(5), 467.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wright, S. (1921). Correlation and causation. *Journal of agricultural research 20*(7), 557–585.

Ziliak, S. T. and D. N. McCloskey (2004). Size matters: the standard error of regressions in the american economic review. *The Journal of Socio-Economics 33*(5), 527–546.

Ziliak, S. T. and D. N. McCloskey (2008). *The cult of statistical significance: How the standard error costs us jobs, justice, and lives*. Ann Arbor, USA: University of Michigan Press.

Online Appendix

A Experiments 2 and 3: how do economists interpret standardized regressions?

Respondents were randomly assigned to two groups: half of them received the results of a standard OLS regression (column 1 of Table A.1), while the other half also received the results of a standardized regression (column 2 of Table A.1). For the sake of comparison, column 3 reports estimates of the absolute economic importance of variables as calculated using the new method expounded in Section 4. This last column was not shown to respondents.[24]

The question was asked as follows, with the information in square brackets only provided to the second group:

"Please find below the results of a simple OLS regression (column 1) [as well as the results of the same regression where all variables were standardized (column 2). Standardization consists of subtracting the mean of the variable and dividing the result by the standard deviation of the variable. This process was done for all variables (y, x_1, x_2, x_3, ...)]. It is assumed that all variables (y, x_1, x_2, ...) are of high interest for economists and policy-makers (e.g. welfare, wages, health, age, etc.). The table only shows results for 5 variables, but 45 supplementary control variables are included in the regression (and also standardized). Please note that the regression is well-specified, and without any statistical bias in the estimated coefficients and standard errors. Coefficients can be interpreted as causal. Variables are normally distributed with different variances and means. Standard errors are reported in parentheses. In percentages, how would you qualify the economic importance of the effect of x_1 on y? And of x_2 on y?"

Results of this experiment are presented in Figure A.1.[25] There is both good

[24]The data generating process of this regression is: $y = 0.1x_1 + 0.3x_2 + 0.2x_3 + 2x_4 - 1x_5 + 1.5x_6 - 0.2x_7 + 0.1x_8 - 0.5x_9 + 0.2x_{10} + \sum_{i=11}^{50} x_i + \epsilon$, where the x_i and ϵ are normally distributed with $\sigma_1 = 140$, $\sigma_2 = 3.3$, $\sigma_3 = 20$, $\sigma_4 = 5.45$, $\sigma_5 = -0.1$, $\sigma_6 = 2.6$, $\sigma_7 = -10$, $\sigma_8 = 1$, $\sigma_9 = -5.8$, $\sigma_{10} = 5$, $\sigma_i = 1 \forall i \in 11...50$, $\sigma_e = 20$.

[25]As explained in appendix B, 10% of responses could not be categorized, and were excluded.

Table A.1 – Results of OLS regressions proposed to the participants of the experiment

	(1) Basic OLS	(2) Standardized OLS	(3) Contribution α_i^2
x_1	0.099***	0.488***	14.3 %
	(0.003)	(0.014)	
x_2	0.324***	0.039***	1.1 %
	(0.118)	(0.014)	
x_3	0.213***	0.153***	4.4 %
	(0.020)	(0.014)	
x_4	2.056***	0.391***	11.9 %
	(0.074)	(0.014)	
x_5	-0.156	-0.001	0.0 %
	(3.928)	(0.014)	
Controls (45 variables)			48.0%
Residuals			20.6%
Sum contributions			100%
Controls (45 variables)	Yes	Yes	Yes
Observations	2583	2583	
R^2	0.504	0.504	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and bad news. On the one hand, a large majority of respondents from the first group (74%) correctly noted that they could not assess the absolute economic importance of variables when only OLS results were available. On the other hand, as much as 35% of respondents from the second group wrongly interpreted the results of the standardized regression in percentage terms, and 24% interpreted the results of the OLS regression in percentage terms. In appendix B, I show that respondents who have completed their PhD are significantly less likely to interpret OLS coefficients in percentage, and significantly more likely to reply that they do not know the answer. Of course, the framing of the question - in percentage - is likely to have influenced some respondents. Despite this caveat, results suggest that a significant proportion of economists might be tempted to interpret the results of OLS and standardized regressions in percentage terms.

The online questionnaire also asked respondents *"What is the sum of regression coefficients in a regression where all variables are standardized (y, x_1, x_2, x_3,*

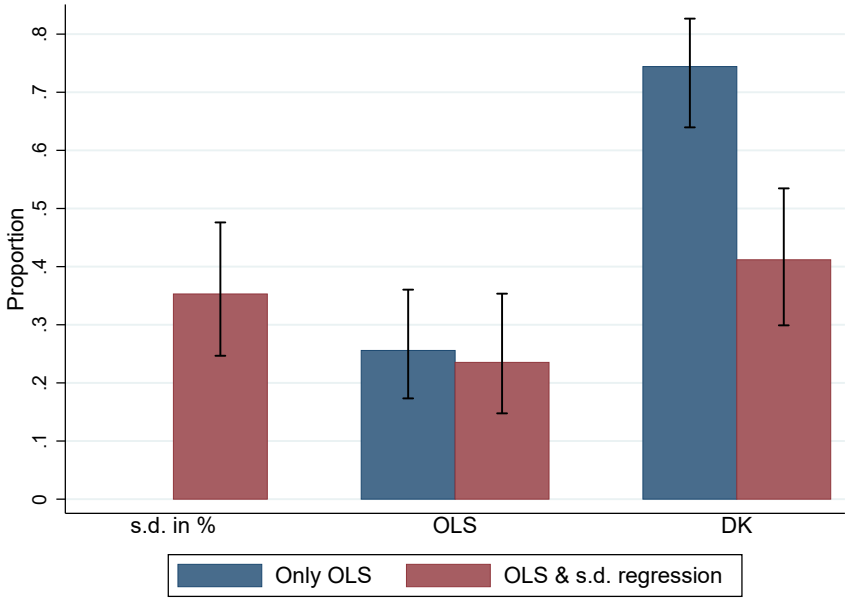


Figure A.1 – Results of experiment 2 (s.d. in percentage: coefficients of standardized regression are interpreted as percentage, OLS: coefficients of OLS regression are interpreted as percentage, DK = do not know).

...)?" The results of this experiment are also mixed (Figure A.2).[26] Only 2% of respondents provided the correct answer. A majority (55%) recognized that they could not answer the question. Worryingly, 34% of respondents answered that the sum is equal to 1, while the rest replied that the sum is equal to 0 or to the r^2. In Appendix B, I show that having a PhD is uncorrelated with the different categories of responses. Respondents' exposure to regression analysis during daily work increases the probability of replying "do not know" and reduces the probability of replying that the sum is equal to 1.

[26] As explained in Appendix B, 1.2% of responses could not be categorized, and were excluded.

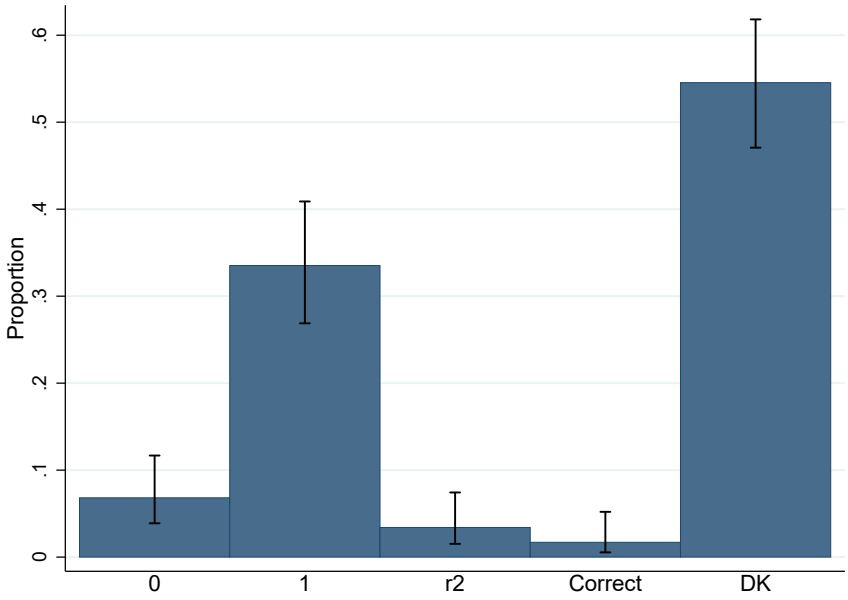


Figure A.2 – Results of experiment 3 (DK = do not know).

B Robustness of experimental results

In this appendix, I show that the results of the three experiments are not much affected by the characteristics of respondents. For each experiment, I first report how answers were categorized, providing a justification for my decision when there is an ambiguity. Then, I use OLS regression to determine whether responses are correlated with "having a PhD", with "exposure to regression analysis" (on a scale from 1 "Never" to 10 "very regularly") and with dummies identifying the different sub-samples. Finally, I analyze how results would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. For this purpose, I use the linear prediction of OLS regressions of respondents' responses on a dummy identifying whether the respondent has a PhD, on exposure to regression analysis (score between 1 and 10), and a dummy identifying the randomization group (for experiments 1 and 2).

B.1 Characteristics of the sample

Table A.2 – Descriptive statistics

Variable	Mean	s.d.	Min.	Max.
Has a PhD	0.6	0.49	0	1
Highest level = PhD student	0.4	0.49	0	1
Highest level = PhD	0.42	0.5	0	1
Highest level = Assistant prof.	0.11	0.32	0	1
Highest level = Associate prof. or prof.	0.06	.24	0	1
Exposed to econometrics	7.70	2.53	1	10
CSAE conference sample	0.57	0.5	0	1
Oxford sample	0.27	0.44	0	1
CSAE, UCLouvain, UNamur sample	0.16	0.37	0	1

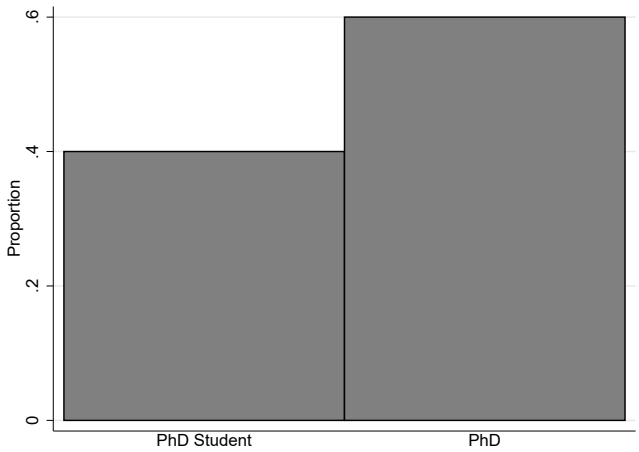


Figure A.3 – Education level/position of respondents

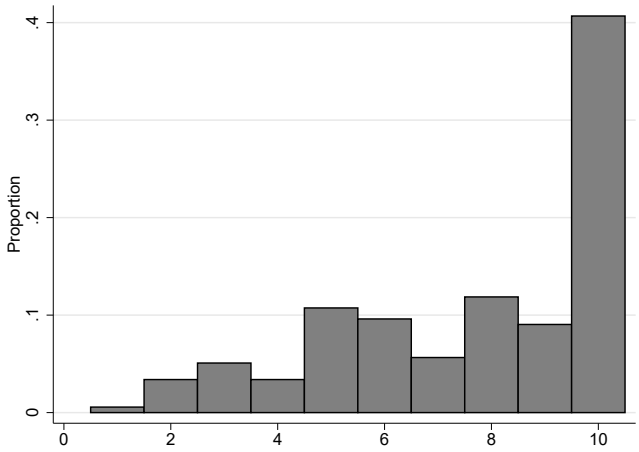


Figure A.4 – Answer to the question: "How often are you exposed to econometrics in your daily work?" (From 1=never to 10=very regularly)

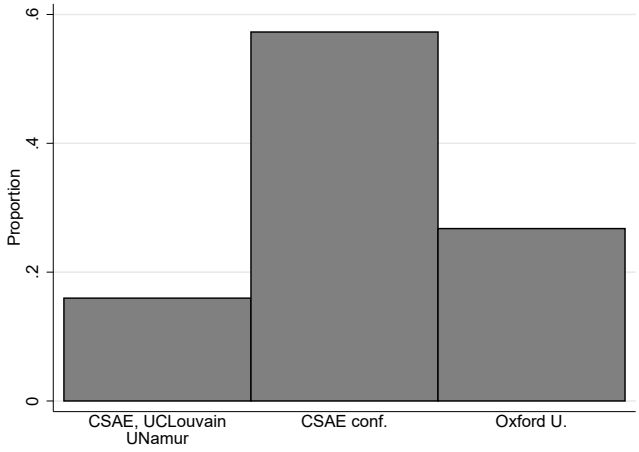


Figure A.5 – Number of respondents from different samples: (1) colleagues from CSAE, UCLouvain and UNamur (pilot survey), (2) Participants to CSAE conferences in 2015 and 2016, and (3) department of economics of Oxford University

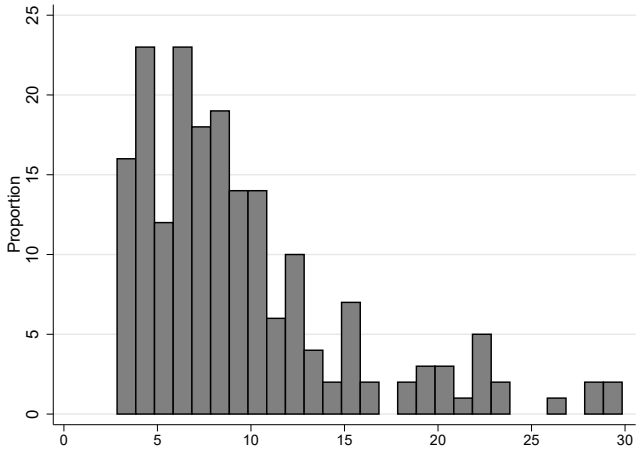


Figure A.6 – Number of minutes taken to reply to the questionnaire

Table A.3 – Balance table

Variable	N	Vector (0,1,5,2) Mean	Vector (0,1,5,2) s.d.	Vector (6,1,5,2) Mean	Vector (6,1,5,2) s.d.	Difference t-test t-stat	Difference t-test p-value
Panel A: experiment 1							
Has a PhD	175	.62	.49	.58	.49	.5	.62
Higher level = PhD student	175	.38	.49	.42	.49	-.5	.62
Higher level = PhD	175	.41	.5	.43	.5	-.26	.79
Higher level = Assistant prof.	175	.12	.33	.11	.33	.33	.74
Higher level = Associate prof. or prof.	175	.08	.27	.04	.27	1.11	.27
Exposed to econometrics	177	7.71	2.49	7.69	2.49	.03	.97
CSAE conference sample	213	.61	.49	.53	.49	1.11	.27
Oxford sample	213	.22	.42	.32	.42	-1.65	.1
CSAE, UCLouvain, UNamur sample	213	.17	.38	.15	.38	.5	.62

Variable	N	OLS regression Mean	OLS regression s.d.	Standardized regression Mean	Standardized regression s.d.	Difference t-test t-stat	Difference t-test p-value
Panel B: experiment 2							
Has a PhD	162	.64	.48	.55	.48	1.13	.26
Higher level = PhD student	162	.36	.48	.45	.48	-1.13	.26
Higher level = PhD	162	.46	.5	.39	.5	.85	.39
Higher level = Assistant prof.	162	.12	.33	.1	.33	.45	.66
Higher level = Associate prof. or prof.	162	.05	.23	.06	.23	-.04	.97
Exposed to econometrics	163	7.96	2.51	7.23	2.51	1.8	.07*
CSAE conference sample	172	.66	.48	.57	.48	1.23	.22
Oxford sample	172	.26	.44	.34	.44	-1.06	.29
CSAE, UCLouvain, UNamur sample	172	.07	.26	.09	.26	-.41	.68

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Experiment 1

The categorization process of responses for the first experiment on absolute importance is explained in table A.7. About 15% of responses could not be categorized and were excluded. Most of these responses argue that the economic importance of the variable x_i is 33% , or is equal to its regression coefficient c_i. Perhaps surprisingly, a few respondents replied that the contributions of variables are 1/14, 5/14 and 2/14 respectively; this corresponds to $|c_i| / (|c_0| + |c_1| + |c_2| + |c_3|)$. This response is close to interpretation (iii), and was classified as such; however, this understanding of economic importance is problematic because it is highly sensitive to a rescaling of the dependent variable (for degree Celsius to Kelvin for example). The categorization of responses for the question on relative importance is described in table A.9. 8% of responses could not be categorized and were therefore excluded.

Table A.4 – Robustness of Experiment 1: absolute economic importance

	(i)		(ii)		(iii)		DK		Not categorized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Randomization group 1	-0.034	-0.039	0.014	0.014	-0.045	-0.042	0.019	0.024	0.045	0.043
	(0.033)	(0.033)	(0.014)	(0.014)	(0.078)	(0.078)	(0.068)	(0.068)	(0.055)	(0.055)
Has a PhD	0.050	0.036	-0.018	-0.020	-0.033	-0.025	0.010	0.022	-0.008	-0.013
	(0.037)	(0.036)	(0.018)	(0.020)	(0.084)	(0.085)	(0.073)	(0.074)	(0.059)	(0.059)
Exposed to econometrics	-0.009	-0.005	-0.002	-0.001	-0.016	-0.018	0.023	0.019	0.004	0.005
	(0.011)	(0.010)	(0.002)	(0.001)	(0.017)	(0.017)	(0.014)	(0.014)	(0.011)	(0.011)
Oxford econ. dep.	0.076	0.080	-0.021	-0.020	-0.014	-0.017	0.016	0.013	-0.058	-0.056
	(0.055)	(0.054)	(0.020)	(0.020)	(0.099)	(0.100)	(0.085)	(0.086)	(0.071)	(0.072)
CSAE/UCLouvain/UNamur	0.009	0.009	-0.013	-0.013	0.039	0.039	0.057	0.057	-0.093	-0.093
	(0.043)	(0.042)	(0.013)	(0.013)	(0.113)	(0.113)	(0.104)	(0.104)	(0.071)	(0.070)
Survey duration (log)		0.076**		0.009		-0.043		-0.068*		0.026
		(0.035)		(0.009)		(0.055)		(0.036)		(0.048)
Constant	0.087	-0.412*	0.032	-0.029	0.704***	0.989**	0.050	0.495*	0.128	-0.044
	(0.087)	(0.237)	(0.032)	(0.033)	(0.155)	(0.391)	(0.129)	(0.294)	(0.114)	(0.326)
Observations	172	172	172	172	172	172	172	172	172	172
R^2	0.039	0.090	0.030	0.037	0.010	0.014	0.020	0.032	0.016	0.018

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are dummies corresponding the categories defined in table A.7. Randomization group 1 is a dummy equal to 0 if the vector of coefficients was (0,1,5,2), and 1 if the vector of coefficients was (6,1,5,2). Exposition to econometrics is measured on a scale from 1 "Never" to 10 "very regularly". The sub-sample from the department of economics of the University of Oxford are identified by the dummy "Oxford econ. dep.". The pilot sub-sample is identified by the dummy "CSAE/UCLouvain/UNamur". The "missing" dummy relates to the participants to the 2015 and 2016 CSAE conferences at the University of Oxford. Survey duration is in second.

	Interpretations						
Raw data	(i)	(ii)	(iii)	DK	Not categorized	Total	Note
1/30 25/30 4/30	5	0	0	0	0	5	
3.33 83.33 13.33	2	0	0	0	0	2	
0.03 0.83 0.13	1	0	0	0	0	1	
1/29 25/29 4/29	1	0	0	0	0	1	The respondent did a computation error
1/30 5/6 2/15	1	0	0	0	0	1	
1/6 5/6 1/3	0	1	0	0	0	1	
1/8 5/8 2/8	0	0	44	0	0	44	
1/8 5/8 1/4	0	0	19	0	0	19	
12.5 62.5 25	0	0	11	0	0	11	
1/14 5/14 2/14	0	0	9	0	0	9	
0.125 0.625 0.25	0	0	4	0	0	4	
1/7 5/7 2/7	0	0	3	0	0	3	The respondent did a computation error
1/10 5/10 2/10	0	0	2	0	0	2	The respondent did a computation error
1/14 5/14 1/7	0	0	2	0	0	2	
1/8 2/8 5/8	0	0	2	0	0	2	The order of the answers is incorrect
1/8 5/8 3/8	0	0	2	0	0	2	The respondent did a typo for the third answer
0.1 approx 0.65 approx	0	0	1	0	0	1	
0.125 0.5 0.25	0	0	1	0	0	1	The respondent did a computation error for the second answer
1/10 1/2 1/5	0	0	1	0	0	1	The respondent did a computation error
1/12 5/12 1/6	0	0	1	0	0	1	This correspond to 1/14, 5/15, 1/7 with a computation error
1/16 5/16 2/16	0	0	1	0	0	1	This correspond to 1/14, 5/15, 1/7 with a computation error
1/6 1/2 1/3	0	0	1	0	0	1	The respondent read 3 instead of 5
12,5 62,5 25	0	0	1	0	0	1	
12.5% 62.5% 25%	0	0	1	0	0	1	
15 50 35	0	0	1	0	0	1	The respondent read 3 instead of 5
15 75 30	0	0	1	0	0	1	The respondent estimated that 1/8 is about 15%, and then muliplied 15% by 5 and 2 to obtain the two other responses
20 50 30	0	0	1	0	0	1	The respondent read 3 instead of 5
20/100 50/100 30/100	0	0	1	0	0	1	The respondent read 3 instead of 5
5/8 2/8 1/8	0	0	1	0	0	1	The order of the answers is incorrect
63 25 12	0	0	1	0	0	1	The order of the answers is incorrect
7 36 14	0	0	1	0	0	1	This correspond to 1/14, 5/15, 1/7
99 99 99	0	0	0	53	0	53	
99 1/2 99	0	0	0	1	0	1	
cannot answer cannot	0	0	0	1	0	1	
100 500 200	0	0	0	0	5	5	
0 0 0	0	0	0	0	3	3	
1 1/5 1/2	0	0	0	0	3	3	
1 5 2	0	0	0	0	3	3	
1/2 1/2 1/2	0	0	0	0	3	3	
1 1 1	0	0	0	0	2	2	
1/3 1/3 1/3	0	0	0	0	2	2	
0.1 0.75 0.15	0	0	0	0	1	1	
1/1 5/1 2/1	0	0	0	0	1	1	
1/100 1/20 1/50	0	0	0	0	1	1	
1/2 3/4 1/2	0	0	0	0	1	1	
1/4 1/2 1/2	0	0	0	0	1	1	
1/5 1/4 11/20	0	0	0	0	1	1	
100 100 100	0	0	0	0	1	1	
25/2 125/2 25	0	0	0	0	1	1	
33 33 33	0	0	0	0	1	1	
33.3333 33.3333 33.33	0	0	0	0	1	1	
4/5 4/5 2/5	0	0	0	0	1	1	
Total	10	1	113	55	32	211	

Figure A.7 – Categorization of responses of Experiment 1: absolute economic importance

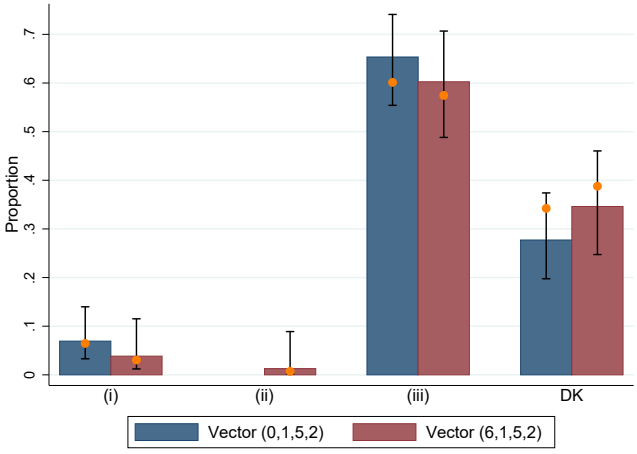


Figure A.8 – Robustness of Experiment 1: absolute economic importance. The orange dots show how the figure would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. It is reassuring to observe that the orange dots are included in the confidence intervals.

	Interpretations					
Raw data	**(i)**	**(ii) and (iii)**	**DK**	**Not categorized**	**Total**	**Note**
25 4	10	0	0	0	10	
5 2	0	119	0	0	119	
1/5 1/2	0	3	0	0	3	Inverse answer
5 1	0	2	0	0	2	Typo for the second answer
5/1 2/1	0	2	0	0	2	
3 2	0	1	0	1	2	Based on absolute importance, one respondent misread c2=3.
2 5	0	1	0	0	1	The order of the answers is incorrect
2.5 2	0	1	0	0	1	x3/x2 and x3/x1
5 1/2	0	1	0	0	1	x1/x2 for the second answer
5 2/5	0	1	0	0	1	x3/x2 for the second answer
5 times 2 times	0	1	0	0	1	
5/1 2/5	0	1	0	0	1	x3/x2 for the second answer
6.1 2.5	0	1	0	0	1	Ratio of estimated absolute importance (0.1;0.65;0.25)
99 99	0	0	45	0	45	
2 99	0	0	1	0	1	
3 99	0	0	1	0	1	
99	0	0	1	0	1	
99 (can't tell, I don't know anything about the units!!) 99	0	0	1	0	1	
cannot answer cannot answer	0	0	1	0	1	
1 1	0	0	0	5	5	
0 0	0	0	0	2	2	
0.5	0	0	0	1	1	
2 3	0	0	0	1	1	
2.5 1.5	0	0	0	1	1	
230 110	0	0	0	1	1	
3 1	0	0	0	1	1	
4 7/4	0	0	0	1	1	
5/6 2/3	0	0	0	1	1	
52 1	0	0	0	1	1	
It depends on what these variables are It depends on what these variables are	0	0	0	1	1	
Total	0	0	0	0	0	

Figure A.9 – Categorization of responses of Experiment 1: relative economic importance

Table A.5 – Robustness of Experiment 1: relative economic importance

	(i)		(ii) and (iii)		DK		Not categorized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Randomization group 1	-0.034	-0.039	0.146**	0.144**	-0.059	-0.053	-0.054	-0.052
	(0.033)	(0.033)	(0.071)	(0.071)	(0.062)	(0.062)	(0.035)	(0.034)
Has a PhD	0.050	0.036	-0.135*	-0.141*	0.108*	0.124*	-0.023	-0.019
	(0.037)	(0.036)	(0.075)	(0.075)	(0.064)	(0.063)	(0.040)	(0.041)
Exposed to econometrics	-0.009	-0.005	-0.013	-0.012	0.038***	0.034***	-0.016**	-0.017**
	(0.011)	(0.010)	(0.015)	(0.015)	(0.011)	(0.011)	(0.007)	(0.008)
Oxford econ. dep.	0.076	0.080	-0.091	-0.089	0.049	0.045	-0.035	-0.036
	(0.055)	(0.054)	(0.094)	(0.095)	(0.080)	(0.082)	(0.047)	(0.046)
CSAE/UCLouvain/UNamur	0.009	0.009	0.020	0.020	0.042	0.042	-0.071**	-0.071**
	(0.043)	(0.042)	(0.103)	(0.103)	(0.095)	(0.095)	(0.029)	(0.029)
Survey duration (log)		0.076**		0.034		-0.087**		-0.023
		(0.035)		(0.048)		(0.036)		(0.019)
Constant	0.087	-0.412*	0.798***	0.575	-0.125	0.448	0.240***	0.389**
	(0.087)	(0.237)	(0.146)	(0.353)	(0.118)	(0.287)	(0.090)	(0.158)
Observations	172	172	172	172	172	172	172	172
R^2	0.039	0.090	0.048	0.050	0.078	0.099	0.056	0.061

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are dummies corresponding the categories defined in table A.9. Randomization group 1 is a dummy equal to 0 if the vector of coefficients was (0,1,5,2), and 1 if the vector of coefficients was (6,1,5,2). Exposition to econometrics is measured on a scale from 1 "Never" to 10 "very regularly". The sub-sample from the department of economics of the University of Oxford are identified by the dummy "Oxford econ. dep.". The pilot sub-sample is identified by the dummy "CSAE/UCLouvain/UNamur". The "missing" dummy relates to the participants to the 2015 and 2016 CSAE conferences at the University of Oxford. Survey duration is in second.

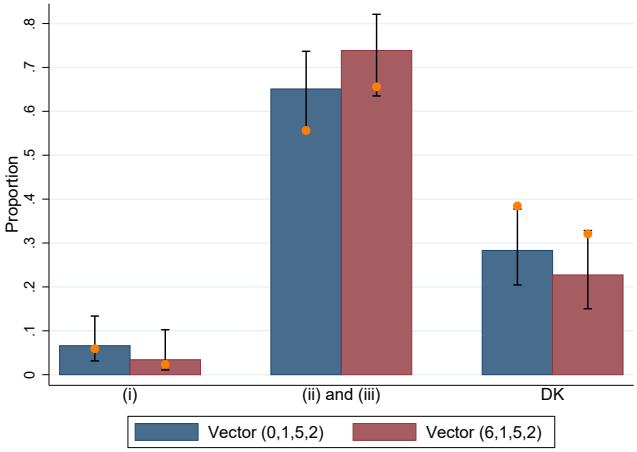


Figure A.10 – Robustness of Experiment 1: relative economic importance. The orange dots show how the figure would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. Five out of six orange dots are included in the confidence intervals. The overall shape of the figure would remain similar with such alternative sample.

B.3 Experiment 2

	Interpretations					
Raw data	**s.d in %**	**OLS**	**DK**	**Not categorized**	**Total**	**Note**
49 4	4	0	0	0	4	
48.8 3.9	3	0	0	0	3	
50 5	3	0	0	0	3	rounding
0.488 0.039	2	0	0	0	2	
46 4	2	0	0	0	2	
50 4	2	0	0	0	2	
23 2	1	0	0	0	1	coef*r2
24 1.52	1	0	0	0	1	coef x r2
25 1	1	0	0	0	1	coef x r2
25 2	1	0	0	0	1	coef x r2
48 3.9	1	0	0	0	1	
50 1	1	0	0	0	1	1 is a typo
70 10	1	0	0	0	1	coef/r2
90 10	1	0	0	0	1	coef/r2
9.9 32.4	0	10	0	0	10	
10 32	0	7	0	0	7	
0.099 0.324	0	3	0	0	3	
0.1 0.3	0	2	0	0	2	
10 30	0	2	0	0	2	
1 3.2	0	1	0	0	1	
1 30	0	1	0	0	1	Typo: 1% instead of 10%
10 40	0	1	0	0	1	
10 99	0	1	0	0	1	
13 44	0	1	0	0	1	coef (1+r2)
20 60	0	1	0	0	1	OLS/r2
3.5 11	0	1	0	0	1	OLS divided by the sum of
4 13	0	1	0	0	1	OLS divided by the sum of coefficients
5 15	0	1	0	0	1	OLS x r2
5 16	0	1	0	0	1	OLS x r2
5 20	0	1	0	0	1	OLS x r2
9 32	0	1	0	0	1	
9.9 32	0	1	0	0	1	
9.99 32.4	0	1	0	0	1	
99 99	0	0	91	0	91	
99 (I don't know the units!!) 99	0	0	1	0	1	
50 50	0	0	0	3	3	
0.6 0.1	0	0	0	1	1	
1/100 1/100	0	0	0	1	1	
1/2 1/2	0	0	0	1	1	
100 10	0	0	0	1	1	
100 100	0	0	0	1	1	
25 75	0	0	0	1	1	
30 3	0	0	0	1	1	t-statistics
4 10	0	0	0	1	1	
4/5 2/5	0	0	0	1	1	
50 75	0	0	0	1	1	
7% 12%	0	0	0	1	1	
70 50	0	0	0	1	1	
8 24	0	0	0	1	1	
precision weight the betas, then multiplying by r^2	0	0	0	1	1	
Total	24	38	92	17	171	

Figure A.11 – Categorization of responses of Experiment 2: absolute economic importance

Table A.6 – Robustness of Experiment 2: absolute economic importance

	s.d. in %		OLS		DK		Not categorized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Randomization group 2	0.315***	0.297***	-0.037	-0.047	-0.299***	-0.269***	0.021	0.019
	(0.057)	(0.056)	(0.067)	(0.070)	(0.076)	(0.078)	(0.049)	(0.049)
Has a PhD	0.014	0.000	-0.194***	-0.203***	0.270***	0.294***	-0.089	-0.091
	(0.054)	(0.055)	(0.073)	(0.073)	(0.079)	(0.078)	(0.058)	(0.058)
Exposed to econometrics	-0.011	-0.008	-0.001	0.001	0.006	0.001	0.007	0.007
	(0.010)	(0.011)	(0.013)	(0.014)	(0.015)	(0.015)	(0.011)	(0.011)
Oxford econ. dep.	-0.058	-0.054	-0.151**	-0.148**	0.149*	0.142*	0.060	0.060
	(0.057)	(0.058)	(0.071)	(0.071)	(0.084)	(0.085)	(0.067)	(0.067)
CSAE/UCLouvain/UNamur	0.081	0.087	-0.163	-0.159	0.108	0.098	-0.026	-0.026
	(0.104)	(0.099)	(0.118)	(0.118)	(0.138)	(0.133)	(0.088)	(0.089)
Survey duration (log)		0.076**		0.047		-0.132***		0.009
		(0.038)		(0.046)		(0.041)		(0.023)
Constant	0.087	-0.406	0.432***	0.132	0.399**	1.250***	0.082	0.024
	(0.088)	(0.271)	(0.130)	(0.330)	(0.154)	(0.315)	(0.104)	(0.186)
Observations	158	158	158	158	158	158	158	158
R^2	0.223	0.248	0.057	0.064	0.164	0.200	0.036	0.036

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are dummies corresponding the categories defined in table A.11. Randomization group 2 is a dummy equal to 0 if the respondent only received the results of a standard OLS regression, and 1 if the respondent also received the results of a standardized regression. Exposition to econometrics is measured on a scale from 1 "Never" to 10 "very regularly". The sub-sample from the department of economics of the University of Oxford are identified by the dummy "Oxford econ. dep.". The pilot sub-sample is identified by the dummy "CSAE/UCLouvain/UNamur". The "missing" dummy relates to the participants to the 2015 and 2016 CSAE conferences at the University of Oxford. Survey duration is in second.

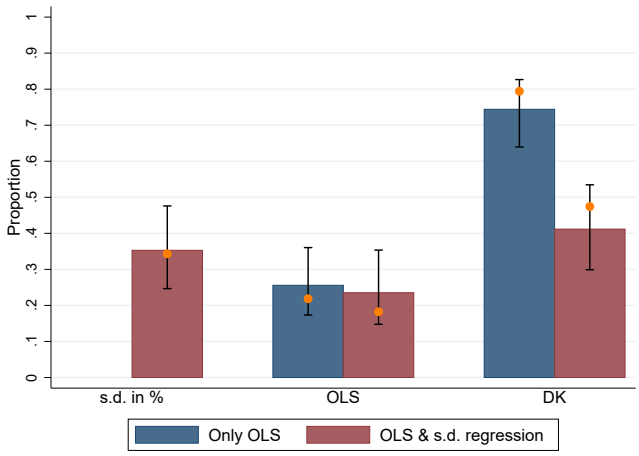


Figure A.12 – Robustness of Experiment 2: absolute economic importance. The orange dots show how the figure would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. The orange dots show how the figure would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. All orange dots are included in the confidence intervals, suggesting that results would be robust to variations in the sampling framework.

	Interpretations					
Raw data	**s.d in %**	**OLS**	**DK**	**Not categorized**	**Total**	**Note**
2.5 0.1	5	0	0	0	5	
2.6 0.08	5	0	0	0	5	
2 0.1	2	0	0	0	2	
2 10	2	0	0	0	2	
2 0.09	1	0	0	0	1	
2 0.5	1	0	0	0	1	Typo for the second answer
2 1/12	1	0	0	0	1	
2 1/13	1	0	0	0	1	
2 1/20	1	0	0	0	1	
2 12	1	0	0	0	1	
2 < 0.5	1	0	0	0	1	
2.4 0.1	1	0	0	0	1	
2.5 0.05	1	0	0	0	1	
2.5 0.07	1	0	0	0	1	
2.5 0.5	1	0	0	0	1	
2.5 12	1	0	0	0	1	
2.5 99	1	0	0	0	1	
2.55 0.08	1	0	0	0	1	
2.6	1	0	0	0	1	
2.67 0.1	1	0	0	0	1	
3 0.02	1	0	0	0	1	Typo for the second answer
3 0.1	1	0	0	0	1	
3 0.12	1	0	0	0	1	
3 1/10	1	0	0	0	1	
300 10	1	0	0	0	1	
40/15 4/49	1	0	0	0	1	
8/3 4/50	1	0	0	0	1	
10 3	0	31	0	0	31	
10 3.2	0	2	0	0	2	
5 3	0	2	0	0	2	ratio OLS*r2 and then ratio OLS
9 3	0	2	0	0	2	
9.7 3.3	0	2	0	0	2	
10 0.001	0	1	0	0	1	
10 3.3	0	1	0	0	1	
21.87 3.27	0	1	0	0	1	b4/b1 and b2/b1
9.6 3.27	0	1	0	0	1	
9.6 3.3	0	1	0	0	1	
9.65 3.27	0	1	0	0	1	
9.7 3.27	0	1	0	0	1	
99 3	0	1	0	0	1	
99 99	0	0	69	0	69	
1 1	0	0	0	4	4	
2 2	0	0	0	4	4	
2 99	0	0	0	3	3	
0.2 0.1	0	0	0	1	1	
0.5 2	0	0	0	1	1	
1 0.1	0	0	0	1	1	
1 1.2	0	0	0	1	1	
20 32	0	0	0	1	1	
200 32	0	0	0	1	1	
6.53 0.006	0	0	0	1	1	
whatever relative magnitude implied by question 3	0	0	0	1	1	
Total	37	47	69	19	172	

Figure A.13 – Categorization of responses of Experiment 2: relative economic importance

Table A.7 – Robustness of Experiment 2: relative economic importance

	s.d. in %		OLS		DK		Not categorized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Randomization group 2	0.399***	0.393***	-0.164**	-0.183**	-0.249***	-0.216***	0.014	0.006
	(0.061)	(0.062)	(0.071)	(0.071)	(0.074)	(0.075)	(0.051)	(0.051)
Has a PhD	-0.046	-0.050	-0.235***	-0.250***	0.257***	0.283***	0.024	0.017
	(0.059)	(0.059)	(0.077)	(0.076)	(0.078)	(0.077)	(0.051)	(0.051)
Exposed to econometrics	0.000	0.001	-0.009	-0.006	0.032**	0.026*	-0.023**	-0.022*
	(0.012)	(0.012)	(0.015)	(0.015)	(0.014)	(0.014)	(0.011)	(0.011)
Oxford econ. dep.	-0.013	-0.012	-0.125	-0.121	0.147*	0.140*	-0.009	-0.007
	(0.066)	(0.066)	(0.079)	(0.079)	(0.083)	(0.084)	(0.061)	(0.061)
CSAE/UCLouvain/UNamur	0.109	0.111	-0.054	-0.047	-0.028	-0.039	-0.027	-0.025
	(0.105)	(0.104)	(0.137)	(0.138)	(0.143)	(0.137)	(0.087)	(0.087)
Survey duration (log)		0.025		0.082*		-0.143***		0.035
		(0.040)		(0.048)		(0.040)		(0.035)
Constant	0.035	-0.128	0.619***	0.091	0.075	0.994***	0.272**	0.044
	(0.103)	(0.282)	(0.141)	(0.350)	(0.141)	(0.300)	(0.111)	(0.260)
Observations	159	159	159	159	159	159	159	159
R^2	0.272	0.274	0.087	0.103	0.178	0.220	0.035	0.041

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are dummies corresponding the categories defined in table A.13. Randomization group 2 is a dummy equal to 0 if the respondent only received the results of a standard OLS regression, and 1 if the respondent also received the results of a standardized regression. Exposition to econometrics is measured on a scale from 1 "Never" to 10 "very regularly". The sub-sample from the department of economics of the University of Oxford are identified by the dummy "Oxford econ. dep.". The pilot sub-sample is identified by the dummy "CSAE/UCLouvain/UNamur". The "missing" dummy relates to the participants to the 2015 and 2016 CSAE conferences at the University of Oxford. Survey duration is in second.

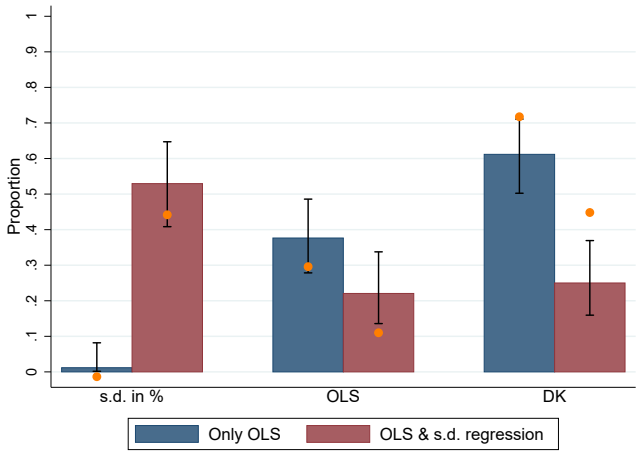


Figure A.14 – Robustness of Experiment 2: relative economic importance. The orange dots show how the figure would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. This figure suggests that the proportion of responses "do not know" would be significantly higher if all respondents had a PhD and were 'very regularly" exposed to regression analysis in their daily work. The overall shape of the figure would remain similar.

B.4 Experiment 3

	Interpretations						
Raw data	**0**	**1**	**r²**	**Correct**	**DK**	**Not categorized**	**Total**
0	11	0	0	0	0	0	11
the unconditional mean of y which should be zero (except if the regression is an AR one)	1	0	0	0	0	0	1
1	0	57	0	0	0	0	57
100	0	2	0	0	0	0	2
1 (Rsquared)	0	0	1	0	0	0	1
R2	0	0	1	0	0	0	1
equal to R, although there are divergence opinion on this because some believe it can be less than R or greater than or equal to R.	0	0	1	0	0	0	1
ssr=tss-sse	0	0	1	0	0	0	1
the R-squared	0	0	1	0	0	0	1
the r2	0	0	1	0	0	0	1
If you square the coefficients on the X you get the R^2. Actual sum is meaningless and can be any value	0	0	0	1	0	0	1
The sum of coefficients multiplied by their standard deviations and all together divided by the standard deviation of y	0	0	0	1	0	0	1
unbounded	0	0	0	1	0	0	1
99	0	0	0	0	96	0	96
b1+b2+b3	0	0	0	0	0	1	1
the sum of the z-scores of the xs	0	0	0	0	0	1	1
Total	12	59	6	3	96	2	178

Figure A.15 – Categorization of responses of Experiment 3

Table A.8 – Robustness of Experiment 3

	0		1		r2		Correct		DK		Not categorized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Has a PhD	-0.014	-0.005	-0.046	-0.059	-0.029	-0.041	0.029	0.024	0.062	0.086	-0.001	-0.004
	(0.045)	(0.045)	(0.078)	(0.078)	(0.032)	(0.033)	(0.018)	(0.016)	(0.082)	(0.083)	(0.017)	(0.018)
Exposed to econometrics	-0.010	-0.012	-0.034**	-0.030*	-0.003	0.000	0.005	0.006	0.037**	0.030*	0.005	0.006
	(0.008)	(0.008)	(0.015)	(0.016)	(0.006)	(0.006)	(0.004)	(0.004)	(0.016)	(0.017)	(0.004)	(0.005)
Oxford econ. dep.	-0.038	-0.040	-0.201**	-0.198**	0.031	0.034	0.021	0.022	0.156*	0.150	0.031	0.032
	(0.046)	(0.045)	(0.084)	(0.084)	(0.040)	(0.040)	(0.034)	(0.034)	(0.094)	(0.095)	(0.025)	(0.026)
CSAE/UCLouvain/UNamur	0.055	0.055	-0.063	-0.063	0.014	0.014	-0.016	-0.016	-0.027	-0.027	0.037	0.037
	(0.070)	(0.070)	(0.105)	(0.105)	(0.040)	(0.040)	(0.013)	(0.013)	(0.114)	(0.113)	(0.040)	(0.040)
Survey duration (log)		-0.048**		0.066		0.065*		0.026		-0.128***		0.018
		(0.024)		(0.050)		(0.036)		(0.026)		(0.046)		(0.014)
Constant	0.152*	0.469**	0.676***	0.237	0.065	-0.364	-0.040	-0.213	0.189	1.033***	-0.043	-0.162
	(0.084)	(0.183)	(0.136)	(0.364)	(0.057)	(0.244)	(0.038)	(0.171)	(0.144)	(0.340)	(0.030)	(0.118)
Observations	171	171	171	171	171	171	171	171	171	171	171	171
R^2	0.020	0.038	0.047	0.057	0.019	0.081	0.023	0.043	0.042	0.074	0.033	0.047

Robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are dummies corresponding the categories defined in table A.15. Exposition to econometrics is measured on a scale from 1 "Never" to 10 "very regularly". The sub-sample from the department of economics of the University of Oxford are identified by the dummy "Oxford econ. dep.". The pilot sub-sample is identified by the dummy "CSAE/UCLouvain/UNamur". The "missing" dummy relates to the participants to the 2015 and 2016 CSAE conferences at the University of Oxford. Survey duration is in second.

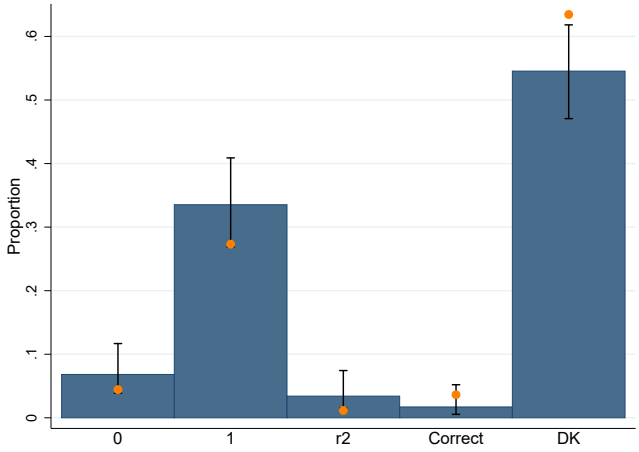


Figure A.16 – Robustness of Experiment 3.
The orange dots show how the figure would look like if the sample was composed of respondents with a PhD who are "very regularly" exposed to regression analysis in their daily work. This figure suggests that the proportion of responses "do not know" would be significantly higher if all respondents had a PhD and were 'very regularly" exposed to regression analysis in their daily work. The overall shape of the figure would remain similar.

C Robustness of application 1

D Robustness of applications 2 and 3

Table A.9 – Application 1: robustness to the addition of an extra control variable x_5

	(1) β	(2) b^*	(3) $\beta_i \frac{\bar{x}_i}{\bar{y}}$	(4) Shapley values	(5) Partial r^2	(6) Semi-partial r^2	(7) New method α_i^2
Panel A - Benchmark results							
x_1	0.997***	0.332***	0.045	0.226	0.398	0.073	0.200
Panel B - x_5 is irrelevant							
x_1	1.006***	0.335***	0.073	0.228	0.402	0.074	0.201
x_5	0.005	0.002	-0.945	0.000	0.000	0.000	0.001
Panel C - x_5 explains 10% of the variance of the error term							
x_1	1.006***	0.335***	0.023	0.228	0.428	0.074	0.191
x_5	1.000***	0.105***	-6.178	0.011	0.100	0.011	0.060
Panel D - x_5 explains 20% of the variance of the error term							
x_1	1.006***	0.335***	0.023	0.228	0.457	0.074	0.188
x_5	1.000***	0.149***	-8.901	0.022	0.200	0.022	0.084
Panel E - x_5 explains 30% of the variance of the error term							
x_1	1.005***	0.335***	0.024	0.227	0.490	0.074	0.187
x_5	1.000***	0.182***	-11.301	0.033	0.300	0.033	0.102
Panel F - x_5 explains 40% of the variance of the error term							
x_1	1.005***	0.335***	0.024	0.227	0.528	0.074	0.186
x_5	1.000***	0.210***	-13.707	0.044	0.400	0.044	0.117
Panel G - x_5 explains 50% of the variance of the error term							
x_1	1.005***	0.335***	0.025	0.227	0.573	0.074	0.185
x_5	1.000***	0.235***	-16.311	0.055	0.500	0.055	0.130
Panel H - x_5 explains 60% of the variance of the error term							
x_1	1.004***	0.334***	0.026	0.227	0.627	0.074	0.186
x_5	1.000***	0.258***	-19.321	0.066	0.600	0.066	0.143
Panel I - x_5 explains 70% of the variance of the error term							
x_1	1.003***	0.334***	0.028	0.227	0.691	0.074	0.186
x_5	1.000***	0.278***	-23.051	0.077	0.700	0.077	0.155
Panel J - x_5 explains 80% of the variance of the error term							
x_1	1.003***	0.334***	0.029	0.227	0.770	0.074	0.188
x_5	1.000***	0.297***	-28.142	0.088	0.800	0.089	0.167
Panel K - x_5 explains 90% of the variance of the error term							
x_1	1.002***	0.334***	0.032	0.227	0.870	0.074	0.190
x_5	1.000***	0.316***	-36.428	0.099	0.900	0.100	0.180

This table illustrates how the statistics associated with x_1 in table 1 change when an extra regressor x_5 is added to the regression model. The different panels vary the importance of the variable x_5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10 – Application 1: robustness to omitted variable bias, mismeasurement bias, reverse causality bias, or misspecification bias

	(1) β	(2) b^*	(3) $\beta_i \frac{\bar{x_i}}{\bar{y}}$	(4) Shapley values	(5) Partial r^2	(6) Semi-partial r^2	(7) New method α_i^2
Panel A - Benchmark results							
x_1	0.997***	0.332***	0.045	0.226	0.398	0.073	0.200
x_2	1.999***	0.664***	-0.076	0.460	0.726	0.294	0.399
x_3	1.003***	0.334***	0.648	0.113	0.502	0.112	0.201
x_4	0.001	0.000	0.001	0.090	0.000	0.000	0.000
Panel B - x_1 is omitted							
x_2	2.336***	0.776***	0.104	0.573	0.711	0.453	0.470
x_3	1.008***	0.336***	-0.039	0.114	0.380	0.113	0.203
x_4	0.334***	0.111***	0.216	0.129	0.048	0.009	0.067
Panel C - x_1 is mismeasured							
$\tilde{x_1}$	0.564***	0.188***	0.025	0.108	0.160	0.030	0.110
x_2	2.200***	0.731***	-0.084	0.518	0.712	0.382	0.426
x_3	1.005***	0.335***	0.649	0.113	0.421	0.112	0.195
x_4	0.202***	0.067***	0.088	0.106	0.021	0.003	0.039
Panel D - x_1 is caused by y							
x_1	2.356***	0.962***	-0.000	0.501	0.750	0.125	0.642
x_2	-0.212***	-0.086***	0.009	0.306	0.022	0.001	0.058
x_3	0.247***	0.101***	0.167	0.089	0.122	0.006	0.067
x_4	0.354***	0.145***	0.162	0.062	0.251	0.014	0.097
Panel E - The DGP is $y = {x_1}^2 + 2x_2 + x_3 + \epsilon$							
x_1	-0.004	-0.002	0.000	0.049	0.000	0.000	0.001
x_2	2.005***	0.706***	0.000	0.401	0.470	0.331	0.436
x_3	1.008***	0.356***	-0.003	0.127	0.253	0.127	0.219
x_4	0.003	0.001	-0.000	0.048	0.000	0.000	0.001

This table illustrates how table 1 changes when the regression is affected by omitted variable bias, mismeasurement bias, reverse causality bias, or misspecification bias. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.11 – Application 1: robustness to multicollinearity in large and small samples

	(1) β	(2) b^*	(3) $\beta_i \frac{\bar{x}_i}{\bar{y}}$	(4) Shapley values	(5) Partial r^2	(6) Semi-partial r^2	(7) New method α_i^2	(8) α_i^2
Panel A - Benchmark results (n=100,000)								
x_1	0.997***	0.332***	0.045	0.226	0.398	0.073	0.200	
x_2	1.999***	0.664***	-0.076	0.460	0.726	0.294	0.399	
x_3	1.003***	0.334***	0.648	0.113	0.502	0.112	0.201	
x_4	0.001	0.000	0.001	0.090	0.000	0.000	0.000	
Panel B - Introducing multicollinearity (n=100,000)								
$\tilde{x_1}$	0.429***	0.143***	0.019	0.144	0.001	0.000	0.086	0.200
$\breve{x_1}$	0.566***	0.189***	0.041	0.144	0.002	0.000	0.114	
x_2	2.000***	0.664***	-0.077	0.416	0.726	0.294	0.399	0.399
x_3	1.003***	0.334***	0.648	0.113	0.502	0.112	0.201	0.201
x_4	0.002	0.001	0.001	0.072	0.000	0.000	0.000	0.000
Panel C - Benchmark results in small sample (n=100)								
x_1	1.061***	0.332***	0.738	0.168	0.437	0.082	0.193	
x_2	1.962***	0.700***	1.174	0.500	0.777	0.370	0.406	
x_3	1.022***	0.328***	-0.196	0.145	0.497	0.105	0.190	
x_4	-0.092	-0.031	-0.025	0.081	0.006	0.001	0.018	
Panel D - Introducing multicollinearity in small sample (n=100)								
$\tilde{x_1}$	1.866	0.582	1.299	0.105	0.015	0.002	0.262	0.194
$\breve{x_1}$	-0.789	-0.247	-0.528	0.104	0.003	0.000	0.111	
x_2	1.961***	0.700***	1.174	0.478	0.779	0.370	0.314	0.404
x_3	1.027***	0.330***	-0.197	0.142	0.502	0.106	0.148	0.190
x_4	-0.106	-0.036	-0.028	0.066	0.008	0.001	0.016	0.021

This table illustrates how table 1 changes when x_1 is measured by two multicollinear variables. In Column (8), the two multicollinear variables are considered together when calculating $\hat{\alpha_1^2}$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.12 – Determinants of migration by dyad: all education levels

	Dependent variable: migration rates, all individuals						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.650***	46.199	45.829	36.056	35.007	20.903	18.572
	(0.027)						
Geo. Dist. (log)	-0.198***	6.911	6.534	5.394	4.991	3.127	2.648
	(0.044)						
Contiguity	-0.295**	0.910	1.865	0.710	1.425	0.412	0.756
	(0.150)						
Genetic Dist.	-0.166**	4.097	3.684	3.198	2.814	1.854	1.493
	(0.068)						
Com. Lang.	0.572***	6.132	6.746	4.785	5.153	2.774	2.734
	(0.089)						
Colonial Link	0.087	0.215	0.494	0.168	0.377	0.097	0.200
	(0.116)						
Pol. restr.	-0.057	0.355	0.577	0.277	0.441	0.160	0.234
	(0.095)						
GDP/cap	2.903***	15.009	13.903	11.713	10.620	6.791	5.634
	(0.639)						
GDP/cap Sq.	-0.183***	0.000	0.000	0.000	0.000	0.000	0.000
	(0.038)						
Educ. Quality	-0.002	0.536	0.489	0.418	0.374	0.243	0.198
	(0.007)						
Import Tariff	-0.004	0.743	0.721	0.580	0.551	0.336	0.292
	(0.008)						
Population (log)	-0.168***	7.685	7.703	5.997	5.884	3.477	3.122
	(0.035)						
Pop 15-24 (% pop.)	-0.000	0.049	0.044	0.038	0.034	0.022	0.018
	(0.021)						
GDP/cap	0.421***	3.301	4.001	2.577	3.056	1.494	1.621
	(0.119)						
Population (log)	0.175***	7.858	7.409	6.132	5.660	3.555	3.003
	(0.026)						
Residuals		0.000	0.000	21.956	23.614	54.754	59.476
R-squared	0.77						
N. of obs	1358.00						
F-stat 1^{st} stage	1974.58						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.13 – Determinants of aspiration rates by dyad: all education levels

	Dependent variable: aspiration rates, all individuals						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.418***	40.600	39.121	28.273	27.462	10.071	10.153
	(0.022)						
Geo. Dist. (log)	-0.121***	5.745	5.275	4.001	3.703	1.425	1.369
	(0.046)						
Contiguity	-0.438***	1.848	3.680	1.287	2.583	0.458	0.955
	(0.116)						
Genetic Dist.	0.134	4.524	3.951	3.151	2.773	1.122	1.025
	(0.088)						
Com. Lang.	0.537***	7.865	8.406	5.477	5.901	1.951	2.181
	(0.101)						
Colonial Link	-0.107	0.361	0.805	0.252	0.565	0.090	0.209
	(0.139)						
Pol. restr.	-0.227**	1.931	3.051	1.345	2.142	0.479	0.792
	(0.092)						
GDP/cap	-0.459	1.846	1.922	1.285	1.349	0.458	0.499
	(0.606)						
GDP/cap Sq.	0.025	0.000	0.000	0.000	0.000	0.000	0.000
	(0.034)						
Educ. Quality	-0.009	3.957	3.509	2.756	2.463	0.982	0.911
	(0.010)						
Import Tariff	-0.002	0.650	0.613	0.453	0.430	0.161	0.159
	(0.010)						
Population (log)	-0.069**	4.318	4.204	3.007	2.951	1.071	1.091
	(0.031)						
Pop 15-24 (% pop.)	0.005	0.865	0.753	0.602	0.529	0.214	0.196
	(0.025)						
GDP/cap	0.488***	5.219	6.144	3.635	4.313	1.295	1.594
	(0.105)						
Population (log)	0.331***	20.270	18.566	14.115	13.032	5.028	4.818
	(0.034)						
Residuals		0.000	0.000	30.363	29.803	75.196	74.048
R-squared	0.62						
N. of obs	1358.00						
F-stat 1^{st} stage	1974.58						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.14 – Determinants of realization rates by dyad: all education levels

	Dependent variable: realization rates, all individuals						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.232***	27.837	28.021	17.707	17.506	4.919	4.494
	(0.029)						
Geo. Dist. (log)	-0.078*	4.572	4.386	2.908	2.740	0.808	0.703
	(0.042)						
Contiguity	0.144	0.750	1.560	0.477	0.974	0.132	0.250
	(0.138)						
Genetic Dist.	-0.300***	12.530	11.431	7.971	7.142	2.214	1.833
	(0.092)						
Com. Lang.	0.035	0.627	0.700	0.399	0.437	0.111	0.112
	(0.106)						
Colonial Link	0.195	0.812	1.889	0.516	1.180	0.143	0.303
	(0.145)						
Pol. restr.	0.170*	1.791	2.957	1.139	1.847	0.317	0.474
	(0.089)						
GDP/cap	3.361***	25.140	23.696	15.992	14.804	4.443	3.800
	(0.761)						
GDP/cap Sq.	-0.208***	0.000	0.000	0.000	0.000	0.000	0.000
	(0.045)						
Educ. Quality	0.007	3.994	3.700	2.541	2.312	0.706	0.593
	(0.010)						
Import Tariff	-0.001	0.452	0.445	0.287	0.278	0.080	0.071
	(0.011)						
Population (log)	-0.099**	7.646	7.777	4.863	4.859	1.351	1.247
	(0.044)						
Pop 15-24 (% pop.)	-0.006	1.153	1.050	0.734	0.656	0.204	0.168
	(0.028)						
GDP/cap	-0.067	0.881	1.084	0.560	0.677	0.156	0.174
	(0.111)						
Population (log)	-0.156***	11.815	11.305	7.516	7.063	2.088	1.813
	(0.037)						
Residuals		0.000	0.000	36.390	37.524	82.328	83.963
R-squared	0.29						
N. of obs	1358.00						
F-stat 1^{st} stage	1974.58						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.15 – Determinants of migration by dyad: less educated individuals

	Dependent variable: migration rates, less educated migrants						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.693***	46.759	46.579	34.869	33.858	18.058	15.707
	(0.032)						
Geo. Dist. (log)	-0.339***	11.429	10.798	8.523	7.849	4.414	3.641
	(0.057)						
Contiguity	-0.399	1.156	2.416	0.862	1.756	0.447	0.815
	(0.246)						
Genetic Dist.	-0.122	2.991	2.691	2.231	1.956	1.155	0.908
	(0.088)						
Com. Lang.	0.276**	2.967	3.219	2.212	2.340	1.146	1.086
	(0.122)						
Colonial Link	-0.230	0.610	1.334	0.455	0.969	0.236	0.450
	(0.193)						
Pol. restr.	-0.057	0.346	0.568	0.258	0.413	0.134	0.191
	(0.115)						
GDP/cap	3.096***	16.521	15.627	12.320	11.359	6.380	5.269
	(0.784)						
GDP/cap Sq.	-0.198***	0.000	0.000	0.000	0.000	0.000	0.000
	(0.047)						
Educ. Quality	-0.011	3.443	3.152	2.567	2.291	1.330	1.063
	(0.008)						
Import Tariff	-0.010	2.117	2.065	1.579	1.501	0.818	0.696
	(0.009)						
Population (log)	-0.159***	7.003	7.030	5.222	5.110	2.705	2.371
	(0.045)						
Pop 15-24 (% pop.)	-0.008	0.964	0.870	0.719	0.633	0.372	0.294
	(0.027)						
GDP/cap	0.077	0.583	0.708	0.435	0.515	0.225	0.239
	(0.167)						
Population (log)	0.071***	3.110	2.942	2.319	2.139	1.201	0.992
	(0.027)						
Residuals		0.000	0.000	25.428	27.312	61.381	66.280
R-squared	0.68						
N. of obs	1198.00						
F-stat 1^{st} stage	1783.23						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.16 – Determinants of aspiration rates by dyad: less educated individuals

	Dependent variable: aspiration rates, less educated migrants						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.422***	38.995	37.824	27.213	26.585	9.504	9.514
	(0.021)						
Geo. Dist. (log)	-0.167***	7.690	7.074	5.366	4.972	1.874	1.780
	(0.050)						
Contiguity	-0.428***	1.698	3.454	1.185	2.427	0.414	0.869
	(0.124)						
Genetic Dist.	0.103	3.451	3.023	2.408	2.125	0.841	0.761
	(0.092)						
Com. Lang.	0.523***	7.696	8.132	5.371	5.716	1.876	2.046
	(0.115)						
Colonial Link	-0.255	0.926	1.970	0.646	1.385	0.226	0.496
	(0.170)						
Pol. restr.	-0.197*	1.647	2.630	1.150	1.848	0.401	0.662
	(0.107)						
GDP/cap	-0.995	3.553	3.601	2.479	2.531	0.866	0.906
	(0.633)						
GDP/cap Sq.	0.055	0.000	0.000	0.000	0.000	0.000	0.000
	(0.035)						
Educ. Quality	-0.017*	7.324	6.529	5.111	4.589	1.785	1.642
	(0.010)						
Import Tariff	-0.000	0.128	0.121	0.089	0.085	0.031	0.030
	(0.011)						
Population (log)	-0.064*	3.857	3.770	2.692	2.650	0.940	0.948
	(0.035)						
Pop 15-24 (% pop.)	-0.008	1.338	1.176	0.934	0.827	0.326	0.296
	(0.027)						
GDP/cap	0.263**	2.714	3.210	1.894	2.256	0.662	0.807
	(0.106)						
Population (log)	0.317***	18.984	17.485	13.248	12.289	4.627	4.398
	(0.035)						
Residuals		0.000	0.000	30.213	29.715	75.628	74.846
R-squared	0.60						
N. of obs	1198.00						
F-stat 1^{st} stage	1783.23						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.17 – Determinants of realization rates by dyad: less educated individuals

	Dependent variable: realization rates, less educated migrants						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.271***	25.993	26.353	16.807	16.724	4.934	4.538
	(0.032)						
Geo. Dist. (log)	-0.173***	8.266	7.948	5.345	5.044	1.569	1.369
	(0.049)						
Contiguity	0.029	0.119	0.253	0.077	0.160	0.023	0.043
	(0.248)						
Genetic Dist.	-0.226**	7.837	7.177	5.068	4.554	1.488	1.236
	(0.104)						
Com. Lang.	-0.247*	3.774	4.168	2.440	2.645	0.716	0.718
	(0.137)						
Colonial Link	0.025	0.094	0.209	0.061	0.132	0.018	0.036
	(0.138)						
Pol. restr.	0.140	1.218	2.033	0.788	1.290	0.231	0.350
	(0.115)						
GDP/cap	4.091***	23.857	23.199	15.426	14.722	4.528	3.995
	(0.920)						
GDP/cap Sq.	-0.254***	0.000	0.000	0.000	0.000	0.000	0.000
	(0.055)						
Educ. Quality	0.006	2.710	2.525	1.752	1.603	0.514	0.435
	(0.010)						
Import Tariff	-0.010	2.878	2.857	1.861	1.813	0.546	0.492
	(0.012)						
Population (log)	-0.095*	5.953	6.081	3.849	3.859	1.130	1.047
	(0.053)						
Pop 15-24 (% pop.)	0.000	0.019	0.017	0.012	0.011	0.004	0.003
	(0.033)						
GDP/cap	-0.185	1.990	2.459	1.286	1.561	0.378	0.423
	(0.147)						
Population (log)	-0.246***	15.292	14.721	9.888	9.342	2.902	2.535
	(0.040)						
Residuals		0.000	0.000	35.340	36.539	81.020	82.781
R-squared	0.28						
N. of obs	1198.00						
F-stat 1^{st} stage	1783.23						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.18 – Determinants of migration by dyad: college graduate individuals

	Dependent variable: migration rates, college graduate migrants						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.648***	36.414	35.424	29.236	28.360	17.983	17.189
	(0.034)						
Geo. Dist. (log)	-0.142***	4.084	3.789	3.279	3.034	2.017	1.839
	(0.049)						
Contiguity	-0.653***	2.236	3.819	1.795	3.057	1.104	1.853
	(0.121)						
Genetic Dist.	0.188*	3.471	3.076	2.787	2.462	1.714	1.492
	(0.111)						
Com. Lang.	0.847***	7.647	8.041	6.140	6.438	3.777	3.902
	(0.128)						
Colonial Link	0.250	0.545	1.168	0.438	0.935	0.269	0.567
	(0.159)						
Pol. restr.	0.118	0.709	1.098	0.569	0.879	0.350	0.533
	(0.088)						
GDP/cap	-0.304	15.730	14.379	12.629	11.512	7.768	6.977
	(0.799)						
GDP/cap Sq.	-0.011	0.000	0.000	0.000	0.000	0.000	0.000
	(0.046)						
Educ. Quality	-0.013	3.165	2.850	2.541	2.282	1.563	1.383
	(0.011)						
Import Tariff	0.025*	4.107	3.890	3.298	3.114	2.028	1.887
	(0.013)						
Population (log)	-0.213***	8.164	8.003	6.555	6.407	4.032	3.883
	(0.043)						
Pop 15-24 (% pop.)	-0.015	1.486	1.315	1.193	1.052	0.734	0.638
	(0.032)						
GDP/cap	1.005***	5.829	6.987	4.680	5.593	2.879	3.390
	(0.124)						
Population (log)	0.190***	6.412	6.162	5.148	4.934	3.167	2.990
	(0.036)						
Residuals		0.000	0.000	19.712	19.941	50.615	51.478
R-squared	0.79						
N. of obs	965.00						
F-stat 1^{st} stage	1385.26						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.19 – Determinants of aspiration rates by dyad: college graduate individuals

	Dependent variable: aspiration rates, college graduate migrants						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.314***	28.954	28.005	21.335	20.739	8.605	8.604
	(0.026)						
Geo. Dist. (log)	-0.091**	4.302	3.969	3.170	2.939	1.279	1.219
	(0.039)						
Contiguity	-0.475***	2.671	4.535	1.968	3.358	0.794	1.393
	(0.137)						
Genetic Dist.	0.298***	9.017	7.945	6.644	5.883	2.680	2.441
	(0.084)						
Com. Lang.	0.467***	6.916	7.230	5.096	5.354	2.055	2.221
	(0.104)						
Colonial Link	0.156	0.559	1.190	0.412	0.881	0.166	0.366
	(0.126)						
Pol. restr.	-0.043	0.430	0.662	0.317	0.490	0.128	0.203
	(0.073)						
GDP/cap	-1.433*	5.625	5.859	4.145	4.339	1.672	1.800
	(0.824)						
GDP/cap Sq.	0.076	0.000	0.000	0.000	0.000	0.000	0.000
	(0.047)						
Educ. Quality	-0.016	6.429	5.755	4.737	4.262	1.911	1.768
	(0.011)						
Import Tariff	0.015	4.120	3.879	3.036	2.873	1.224	1.192
	(0.011)						
Population (log)	-0.084**	5.259	5.125	3.875	3.795	1.563	1.575
	(0.042)						
Pop 15-24 (% pop.)	0.046	7.338	6.453	5.407	4.779	2.181	1.982
	(0.033)						
GDP/cap	0.814***	7.747	9.233	5.709	6.837	2.303	2.837
	(0.106)						
Population (log)	0.192***	10.634	10.161	7.836	7.525	3.161	3.122
	(0.033)						
Residuals		0.000	0.000	26.314	25.944	70.280	69.277
R-squared	0.63						
N. of obs	965.00						
F-stat 1^{st} stage	1385.26						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.20 – Determinants of realization rates by dyad: college graduate individuals

	Dependent variable: realization rates, college graduate migrants						
	β	α_i^1		α_i^2		α_i^3	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network (% pop., in log)	0.334***	33.853	33.961	23.913	23.514	15.338	14.163
	(0.029)						
Geo. Dist. (log)	-0.051	2.637	2.524	1.863	1.747	1.195	1.052
	(0.043)						
Contiguity	-0.178	1.098	1.934	0.776	1.339	0.498	0.807
	(0.124)						
Genetic Dist.	-0.110	3.649	3.335	2.578	2.309	1.653	1.391
	(0.100)						
Com. Lang.	0.380***	6.192	6.714	4.374	4.649	2.805	2.800
	(0.101)						
Colonial Link	0.094	0.369	0.815	0.261	0.564	0.167	0.340
	(0.135)						
Pol. restr.	0.161**	1.751	2.796	1.237	1.936	0.793	1.166
	(0.070)						
GDP/cap	1.129	24.407	22.418	17.240	15.522	11.058	9.349
	(0.836)						
GDP/cap Sq.	-0.086*	0.000	0.000	0.000	0.000	0.000	0.000
	(0.049)						
Educ. Quality	0.003	1.356	1.259	0.958	0.872	0.614	0.525
	(0.011)						
Import Tariff	0.010	2.880	2.812	2.034	1.947	1.305	1.173
	(0.012)						
Population (log)	-0.129***	8.944	9.041	6.318	6.260	4.053	3.771
	(0.044)						
Pop 15-24 (% pop.)	-0.062**	10.743	9.799	7.589	6.784	4.868	4.086
	(0.030)						
GDP/cap	0.191	1.999	2.470	1.412	1.710	0.906	1.030
	(0.131)						
Population (log)	-0.002	0.122	0.121	0.086	0.084	0.055	0.050
	(0.033)						
Residuals		0.000	0.000	29.363	30.762	54.692	58.297
R-squared	0.48						
N. of obs	965.00						
F-stat 1^{st} stage	1385.26						

Robust standard errors, clustered by country of origin, in parentheses. The network variable is instrumented with its 10-year lag. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.21 – Determinants of women's total number of births: OLS regression

	OLS regression						
	β	$\alpha^1 i$		$\alpha^2 i$		$\alpha^3 i$	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(1+density)	-0.045***	2.552	2.610	1.814	1.778	0.699	0.579
	(0.002)	[2.4-2.7]	[2.5-2.8]	[1.7-1.9]	[1.7-1.9]	[0.7-0.7]	[0.5-0.6]
Caloric suitability index	0.009***	0.895	0.973	0.636	0.663	0.245	0.216
	(0.001)	[0.7-1.1]	[0.8-1.2]	[0.5-0.8]	[0.5-0.8]	[0.2-0.3]	[0.2-0.3]
Distance to water	0.025***	1.130	1.248	0.803	0.850	0.309	0.277
	(0.003)	[1-1.3]	[1.1-1.4]	[0.7-0.9]	[0.7-1]	[0.3-0.4]	[0.2-0.3]
Married	0.924***	11.032	10.852	7.840	7.392	3.022	2.409
	(0.009)	[10.9-11.2]	[10.7-11]	[7.7-8]	[7.3-7.5]	[3-3.1]	[2.4-2.5]
Mean marriage	-0.174***	0.870	0.865	0.618	0.590	0.238	0.192
	(0.032)	[0.7-1.1]	[0.7-1.1]	[0.5-0.8]	[0.5-0.8]	[0.2-0.3]	[0.1-0.2]
Child mortality	1.961***	7.867	9.645	5.591	6.570	2.155	2.141
	(0.024)	[7.7-8]	[9.5-9.9]	[5.5-5.7]	[6.4-6.7]	[2.1-2.2]	[2.1-2.2]
Mean child mortality	0.608***	1.068	1.134	0.759	0.772	0.292	0.252
	(0.086)	[0.8-1.2]	[0.9-1.3]	[0.6-0.9]	[0.6-0.9]	[0.2-0.3]	[0.2-0.3]
ln(GDP per capita)	-0.047***	1.616	1.647	1.148	1.122	0.443	0.366
	(0.005)	[1.4-1.8]	[1.4-1.9]	[1-1.3]	[0.9-1.3]	[0.4-0.5]	[0.3-0.4]
Woman's education	-0.019***	10.168	10.654	7.226	7.257	2.785	2.365
	(0.002)	[10-10.4]	[10.5-10.8]	[7.1-7.4]	[7.2-7.4]	[2.7-2.8]	[2.3-2.4]
Woman's education squared	-0.005***						
	(0.000)						
Education in cluster	-0.041***	6.554	6.554	4.658	4.465	1.795	1.455
	(0.005)	[6.3-6.8]	[6.3-6.8]	[4.5-4.8]	[4.3-4.6]	[1.7-1.9]	[1.4-1.5]
Education in cluster squared	-0.002***						
	(0.000)						
Age dummies	Yes	46.558	43.615	33.089	29.710	12.752	9.682
		[46.2-46.9]	[43.2-44]	[32.9-33.3]	[29.5-29.9]	[12.7-12.8]	[9.6-9.7]
Country dummies	Yes	9.690	10.205	6.887	6.951	2.654	2.265
		[9.6-9.8]	[10.1-10.4]	[6.8-7]	[6.9-7]	[2.6-2.7]	[2.2-2.3]
Residuals				28.929	31.881	72.610	77.800
				[28.8-29.1]	[31.7-32]	[72.4-72.9]	[77.6-78]
Observations	490669						
R^2	0.605						

Robust standard errors, clustered at the cluster level, in parentheses. Bootstrap confidence intervals (95%) in square brackets. All specifications include age dummies and country fixed effects. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.22 – Determinants of women's total number of births: IV regression

		IV regression					
	β	$\alpha^1 i$		$\alpha^2 i$		$\alpha^3 i$	
		$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i	$\hat{\delta}_i$	s_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(1+density)	-0.158***	8.849	9.024	6.309	6.182	2.575	2.162
	(0.015)	[8-9.6]	[8.2-9.8]	[5.7-6.9]	[5.6-6.7]	[2.3-2.8]	[1.9-2.4]
Caloric suitability index	0.011***	1.033	1.119	0.736	0.767	0.301	0.268
	(0.002)	[0.8-1.2]	[0.9-1.3]	[0.6-0.9]	[0.6-0.9]	[0.2-0.4]	[0.2-0.3]
Distance to water	-0.017**	0.730	0.804	0.520	0.551	0.212	0.192
	(0.007)	[0.5-1]	[0.5-1.1]	[0.3-0.7]	[0.3-0.8]	[0.1-0.3]	[0.1-0.3]
Married	0.921***	10.770	10.563	7.678	7.237	3.134	2.530
	(0.009)	[10.6-11]	[10.4-10.7]	[7.5-7.8]	[7.1-7.4]	[3.1-3.3]	[2.5-2.7]
Mean marriage	-0.375***	1.840	1.826	1.312	1.251	0.536	0.437
	(0.043)	[1.6-2.1]	[1.6-2.1]	[1.1-1.5]	[1.1-1.4]	[0.5-0.6]	[0.4-0.5]
Child mortality	1.960***	7.704	9.417	5.493	6.452	2.242	2.256
	(0.024)	[7.5-7.9]	[9.2-9.6]	[5.4-5.6]	[6.3-6.6]	[2.2-2.4]	[2.2-2.4]
Mean child mortality	0.602***	1.036	1.096	0.738	0.751	0.301	0.263
	(0.090)	[0.8-1.3]	[0.9-1.4]	[0.6-0.9]	[0.6-0.9]	[0.2-0.4]	[0.2-0.3]
ln(GDP per capita)	-0.006	0.194	0.197	0.139	0.135	0.057	0.047
	(0.008)	[0-0.5]	[0-0.5]	[0-0.4]	[0-0.4]	[0-0.2]	[0-0.1]
Woman's education	-0.020***	9.983	10.414	7.117	7.135	2.905	2.495
	(0.002)	[9.7-10.3]	[10.2-10.7]	[6.9-7.3]	[7-7.3]	[2.8-3.1]	[2.5-2.6]
Woman's education squared	-0.005***						
	(0.000)						
Education in cluster	0.001	2.688	2.786	1.916	1.909	0.782	0.667
	(0.008)	[9.7-10.3]	[10.2-10.7]	[6.9-7.3]	[7-7.3]	[2.8-3.1]	[2.5-2.6]
Education in cluster squared	-0.002***						
	(0.000)						
Age dummies	Yes	45.615	42.612	32.521	29.194	13.275	10.208
		[45.2-46.1]	[42.2-43.1]	[32.3-32.8]	[29-29.5]	[13.2-13.9]	[10.1-10.7]
Country dummies	Yes	9.559	10.143	6.815	6.950	2.782	2.430
		[9.3-9.7]	[10-10.3]	[6.7-6.9]	[6.8-7.1]	[2.7-2.9]	[2.4-2.5]
Residuals				28.704	31.487	70.897	76.045
				[28.5-28.9]	[31.3-31.7]	[69.7-71.1]	[75-76.3]
Observations	490669						
R^2	0.601						

Robust standard errors, clustered at the cluster level, in parentheses. Bootstrap confidence intervals (95%) in square brackets. All specifications include age dummies and country fixed effects. Dispersion $d(x_i)$ is measured by the mean absolute deviation around the mean in Columns (2), (4), and (6), and by standard deviation in Columns (3), (5) and to (7). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$