Presented at the DLSU Research Congress 2015
De La Salle University, Manila, Philippines
March 2-4, 2015

Evidence for causal interpretations of the Easterlin Paradox: Propensity Score Matching and Dose-Response Functions on Philippine data from the 2004 and 2014 World Values Surveys

Gerardo L. Largoza[1,*], Immy Felyss C. Favorada[1], Czarina Viele G. Reinante[1], Stephanie Grace T. Tan[1] and Jemimah S. Thai[1]
[1]*School of Economics, DLSU*
**Corresponding Author: gerardo.largoza@dlsu.edu.ph*

Abstract: We build on previous work (Largoza, 2014) investigating whether Easterlin's controversial yet correlational claims of a non-monotonic relationship between income and happiness (1974, 2010) are amenable to causal interpretation. This time, we use Philippine data from Waves 4 and 6 of the World Values Survey (released 2004 and April 2014) that feature multiple measures of relative income and self-reported well-being, and subject these to a weak and strong test of causality. In the weak test, we run Propensity Score Matching techniques by generating propensity scores based on exogenous socio-demographic characteristics and match individuals based on score similarity; this, in effect, mimics random assignment into treatment and control groups for non-experimental datasets, and allows us to conclude whether higher incomes "cause" individuals to be happier than their low-income counterparts, as though some experiment had been carried out. In the strong test, we construct dose-response functions that treat levels of income more incrementally as "doses" (as opposed to simpler high-low income distinctions); this allows us to tease out possible non-linearities in the happiness-income relationship to see over which levels the relationship changes, again as if an experiment "dosing" randomly selected groups with various levels of income had been carried out. As far as we are aware, this is the first use of techniques from the emerging causal effects literature to re-examine Easterlin's claims about income and happiness.

Key Words: Easterlin Paradox; happiness economics; propensity score matching; dose-response models; causal effects

1. INTRODUCTION

The "Easterlin Paradox", the claim that once certain needs are met, increased incomes no longer correlate with increased happiness, remains controversial forty years after it was first proposed. Despite rigorous and nuanced debate in academic journals (see Easterlin 1974, 2010, but also Stevenson & Wolfers, 2008; Kahneman & Deaton, 2010), the soundbite-friendly claim has, in the Philippines, spawned triumphalist headlines and a discourse toward poverty and inequality that threatens to become even more glib[1].

This paper is motivated by two questions: first, in the Philippines, is there any evidence that income and happiness are indeed uncorrelated, such that perhaps wealthy households are not much happier than poor households? Second, can we subject these data to methods that will support a causal, not merely correlational, interpretation of

[1] "Filipinos 'third happiest' in Asia (National Competitiveness Council Philippines, 2013), "Survey after survey, the Philippines ranks among the happiest countries in the world" (Flifestyle.com, 2013), "Filipino millennials happiest, least stressed in the world" (Inquirer.net, 2013)

the link between income and happiness; that is, can we show that higher incomes *cause* households to become less (or more) happy?

We highlight two contributions to the growing literature on "happiness economics". First, the use of relative (rather than absolute) income when testing for the paradox. Relative income, or the degree to which one feels wealthier or poorer compared to a reference group, was Easterlin's preferred measure even in the seminal 1974 paper, as it suggested a mechanism by which social comparisons and status anxiety might make individuals less happy despite becoming wealthier. Data on relative income are rarely collected in standard economic surveys, but proxy measures fortunately appear in the World Values Survey.

Second, the use of methods from the causal effects literature to see if the relationship between income and happiness is amenable to causal interpretation. To do so under the so-called Neyman-Rubin causal model (2005, 1974), we would need to know the counterfactual outcome – how happy a poor household would have been had it not been poor. In a laboratory setting, this counterfactual might be observed by randomly assigning households to become "wealthy" or "poor", but since no such experiment was carried out on the WVS respondents, we would need to turn to quasi-experimental methods meant to mimic random assignment. In this paper, we implement two, and they provide both a "weak" test of the Easterlin paradox, and a "strong" test.

Propensity Score Matching (PSM) takes non-experimental data and allows us to generate "propensity scores" for respondents based on a series of exogenous covariates. We can then select from a variety of matching algorithms to pair respondents with similar propensity scores. So long as the criteria of conditional independence and a large enough region of common support are fulfilled, this in effect mimics random assignment. The matched pairs would be as similar to each other as possible, save for the fact that one would be considered "treated" (in this case, wealthy) and the other considered "untreated/control" (i.e., poor). A t-test run on the difference in happiness between the two would not just show the impact of income on well-being; it would in addition be interpretable as the causal effect of income on well-being.

Nevertheless, we consider PSM a "weak" test because it restricts treatment to a binary variable, wealthy or poor. A "stronger" test would involve estimating a "dose-response" function, allowing for more incremental "doses" of income and measuring their causal effect on well-being. To do this, we similarly mimic the creation of treatment and control groups, but this time generate propensity scores for every level of income, and carry out tests to ensure a balanced distribution of propensity scores across income levels. We then choose an appropriate "link function" to describe the relationship between the treatment variable and its covariates, and on this basis, implement Stata's *glmdose* command to estimate the model and graph dose-response and treatment effect functions. These functions will reveal any underlying non-linearities in the relationship between income and well-being, perhaps showing at what level of income the impact on well-being begins to diminish.

2. DATA AND METHODS

Data. We use the two most recent datasets on happiness and life satisfaction available for the Philippines. These were obtained from Waves 4 and 6 of the World Values Survey (WVS), released in 2004 and April 2014 respectively, but for which households were actually interviewed in 2001 and 2012. Each survey collected 1,200 responses from voting-age adults, divided uniformly across four major geographical areas: NCR, Luzon, Visayas and Mindanao.

The WVS features two measures of subjective well-being: feelings of happiness and life satisfaction. The former refers to current positive affect while the latter to a longer-term sense of fulfillment. For happiness, the question asked is *"Taking all things together, would you say you are…"*, measured on a four-point Likert scale that we re-oriented for our purposes to run from "Not at all happy" to "Very happy". For life satisfaction, the question is *"All things considered, how satisfied are you with your life as a whole these days?"*, measured this time on a ten-point Likert scale from "Completely dissatisfied" to "Completely satisfied". We drop observations whose responses were indeterminate.

Measuring income is less straightforward, as the WVS does not ask respondents to state their incomes directly. We exploit three proxies instead: the *presence or absence of household savings*, *self-reported social class*, and *scale of income*. In each case, for the purpose of the "weak" test, we transform each income proxy into a binary variable: for household savings, 1=with savings and 0=no savings, for social class, 1=upper/upper middle/lower middle class and 0=working/lower class; and for scale of income, 1=income levels 8 to 10 and 0=income levels 1-7. For the "strong" test, we allow the proxies to take on their Likert scale forms.

PSM implementation: Choosing covariates. The first step in the propensity score matching technique is formulating the propensity score model. For this, we use five covariates: (1) gender, (2) age, (3) region, (4) marital status, and (5) educational attainment. These covariates were chosen because their combination yielded a 0.00 p-value and around 6% to 12% pseudo-R^2 for most of the matching algorithms. Moreover, the effect and significance of the variables in this combination remained consistent despite our adding more variables to the initial set.

The final propensity score model we use is:

$$happy=\beta_1+\beta_2 income+\beta_3 gender+\beta_4 age+\beta_5 region+\beta_6 marital+\beta_7 educ+u \quad \text{(Eq. 1)}$$

where income = 1 if the particiant is non-poor, 0 otherwise
gender = 1 if the participant is male, 0 otherwise
region = 1 if the participant is from NCR, 0 otherwise
marital = 1 if the participant is living with a partner, 0 otherwise.

This participation equation measures the happiness level of the individuals sampled given their income level, gender, age, region of origin, marital status and educational attainment. This equation is used to estimate the treatment effect of being non-poor on an individual's happiness level.

Selecting a matching algorithm. We use Stata's *pstest* command which calculates several measures of variable balancing before and after matching. The balance is checked by this command by considering the t-tests for equality of means in the treated and non-treated groups before and after matching. Then, to be considered as good balancing, these t-values should be insignificant after matching. After using *pstest*, the results show the best matching algorithm to use is a bi-weight or quartic kernel matching algorithm.

Running t-tests. After establishing the propensity score model and choosing the matching algorithm, we then run t-tests. This calculates the "program treatment effect" which shows the difference in the outcome of the treatment and control groups. The following Stata 13 commands implement propensity score matching through the kernel bi-weight algorithm:

For Happiness:

```
psmatch2 savings gender age region marital educ, out(happy) logit ate kernel k(biweight)
psmatch2 class gender age region marital educ, out(happy) logit ate kernel k(biweight)
psmatch2 income gender age region marital educ, out(happy) logit ate kernel k(biweight)
```

For Life Satisfaction:

```
psmatch2 savings gender age region marital educ, out(life) logit ate kernel k(biweight)
psmatch2 class gender age region marital educ, out(life) logit ate kernel k(biweight)
psmatch2 income gender age region marital educ, out(life) logit ate kernel k(biweight)
```

Aside from calculating the average treatment effect (ATE), these commands also provide the average treatment effect on the treated (ATT, based on the counterfactual of those treated) and allow common support graphing and covariate imbalance testing.

Dose-response functions. In testing the "strong" version of the Easterlin paradox, we also undertake the three major steps involved in PSM, to which are added the following procedures:

Estimating Generalized Propensity Scores (GPS). We use a generalised linear model (GLM) to generate the propensity scores of respondents and to conduct the balancing tests to assure that the distribution of the GPS is properly balanced. This method mimics a normal distribution on a non-normal distribution of treatments to allow the analysis of treatments that do not have a normal distribution (Guardabascio & Ventura, 2014). In our case, the income variables: savings, social class and scale of income are determined to have a Gamma distribution.

Choosing a link function. The link function determines how the mean is related to the covariates, and is determined depending on the relationship of the treatment variable and its covariates (Guardabascio & Ventura, 2014). This means that the link function allows the treatment variable to be a part of an exponential family, in our case the Gamma distribution, letting the variance vary with the mean and the covariates. Thus, for us to determine the appropriate link function for the Gamma distribution, we referred to the Stata (n.d.) manual on the generalised structure equation model family and link options. From this manual, we learn that the logarithmic function is the appropriate link function for our exponential family, Gamma.

Implementing the dose-response model. After determining the GPS and choosing the appropriate link function, we then implement the dose-response model. This dose-response model measures the effects of an increase in the relative income of the respondent to one's level of happiness. However, due to the limited scale (1-4) of the level of happiness variable, Stata only generates generalized propensity scores for life satisfaction. Stata's *glmdose* command both computes generalised propensity scores and executes the dose-response model. This command also allows for the balancing tests and the drawing of graphs of both the dose-response function and the treatment effect function.

After implementing both the propensity score matching technique and dose-response model for the "weak" and "strong" versions of the Easterlin Paradox, respectively, we summarise results in the following section.

3. RESULTS AND DISCUSSION

Table 1 below shows results from the "weak" test (PSM) conducted using a bi-weight kernel matching technique; we record positive ATE and ATT values on all combinations of income and happiness variables for both waves. These results indicate a significant positive relationship between income and happiness, and that this relationship may be interpreted causally.

Table 1. Results of propensity score matching (average treatment effects and average treatment effects on the treated) for various income proxies on measures of well-being

OUTCOME	TREATMENT	YEAR	ATT	ATE	*t*
Feeling of happiness	SAVINGS	2001	0.18 ***	0.18 ***	3.31
		2012	0.18 ***	0.12 ***	2.74
	SOCIAL CLASS	2001	0.10 ***	0.13 ***	1.91
		2012	0.06 ***	0.03 ***	1.15
	SCALE OF INCOME	2001	0.20 ***	0.18 ***	3.13
		2012	0.09 ***	0.08 ***	1.35
Life satisfaction	SAVINGS	2001	0.72 ***	0.80 ***	3.37
		2012	0.33 ***	0.14 ***	1.25
	SOCIAL CLASS	2001	0.70 ***	0.75 ***	3.77
		2012	0.27 ***	0.24 ***	1.43
	SCALE OF INCOME	2001	1.15 ***	1.12 ***	5.22
		2012	0.78 ***	0.86 ***	2.98

Source: Authors' calculations

From Table 2 below, savings, social class and scale of income appear to have positive causal effects on life satisfaction, except for the 2001 result for savings. Despite these results being mostly insignificant, the 2001 results show significant evidence that social class and scale of income cause a 0.6464 and 0.3563 positive impact on life satisfaction, respectively. Also, we note decisive evidence against the balancing property; even though it was satisfied at levels lower than 0.01.

Furthermore, Figures 1 and 2 show that the graphs of the causal relationship between life satisfaction and either savings, social class and scale of income are all non-linear, the positive effect wavering at a certain point, and then it either diminishing or remaining at that point.

Table 2. Results of dose-response regressions of Life Satisfaction against income proxies

OUTCOME	TREATMENT	YEAR	COEFF	P-VAL
Life Satisfaction	SAVINGS	2001	-0.1599664	0.812
		2012	0.450622	0.492
	SOCIAL CLASS	2001	0.646004	0.003
		2012	0.4000407	0.367
	SCALE OF	2001	0.3562694	0.009

INCOME	2012	0.0153104	0.884

BALANCING PROPERTY

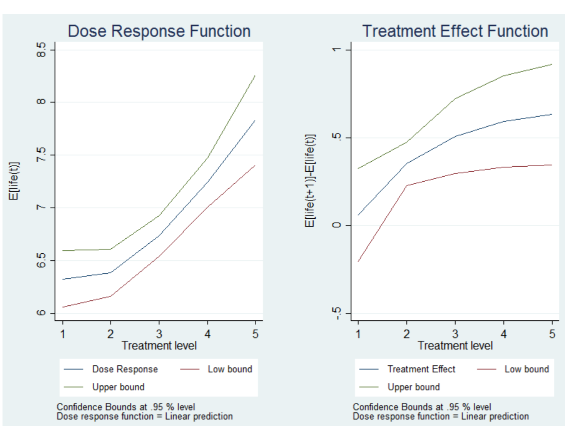
For all outcomes, there is decisive evidence against the balancing property. The balancing property is satisfied at a level lower than 0.01.

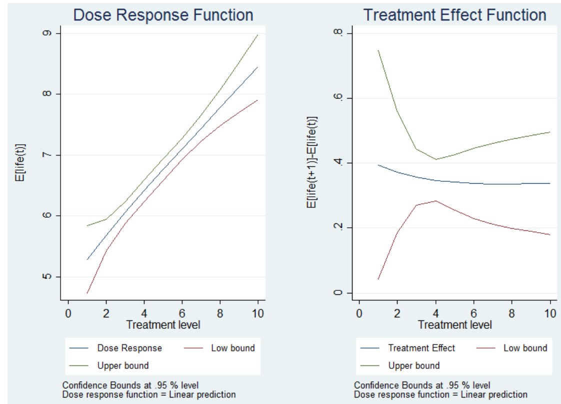
Source: Authors' calculations

Based on the results above, the "weak" test of the Easterlin paradox shows a positive causal relationship between relative income and happiness. After establishing the existence of such causal relationship, the "strong" version of the paradox shows that the positive causal relationship between relative income and happiness is non-linear, and that the relationship either tapers off or stays the same after reaching a certain income level.

From the figures below we notice that the dose response function is still increasing even past income level 10. We suppose that the relationship will taper off at a much higher level than the highest measure of relative income in the Philippines. In other words, an average Filipino's perception of her highest level of income may well be lower compared to that of people from other countries. This means that if an average Filipino does not expect to earn, USD 20,000 or PHP 1 million as her annual income, it implies that the level at which it will taper off is likely higher than level 10, perhaps as high as level 20, 30 or more. We can also observe that the slopes of the graphs are changing; this suggests that the data gathered from WVS can only confirm the beginning of the paradox.

Figs 1 and 2. Graphs of dose-response functions for 2001 social class and scale of income respectively





Besides the PSM and dose response tests, we also conduct a naïve regression using this equation for both datasets:

$$happy = income + income^2 + u_i \quad \text{(Eq. 4.1)}$$

The regression results indicate support for the Easterlin paradox: that income does not strongly increase happiness in the Philippines. Moreover, the change in the sign of the coefficient for *income²* means that there is no continuous positive relationship between income and happiness. Instead, the relationship is like a downward opening parabola (concave), and the peak of this parabola is the maximum value for *income²*. Comparing the naïve regression results with that of PSM and dose response, we can see that the naïve results are consistent with others. This is because the concavity of the naïve results is consistent with the dose response results which show the non-linear relationship between income and happiness. Moreover, in comparing the results of the naïve regression with that of the ATT results we got from PSM, it is evident that our results are not only statistically significant, but are also economically significant. This is because the ATT results are relatively higher than those of the naïve results. This implies that the effect of income on happiness for Filipino individuals is not just a minimal effect, but is a significant one.

4. CONCLUSIONS

Our results show that contrary to the most controversial claims of the Easterlin paradox, there exists a significant positive relationship between income and happiness on both sampled years, and that these results may be interpreted causally. Despite news articles and surveys that report on Filipinos being the happiest people on earth (Lopez, 2012); and Filipino anthropologists such as Michael Tan (2005), describing Filipinos as “so poor, so happy.” our results say otherwise: money does make Filipinos happier.

Aside from this, although most of our results for the “strong” version of the Easterlin Paradox appear to be insignificant, there is still some evidence from the 2001 results that show how Philippine data exhibits a positive relationship between income and happiness for Filipino individuals only up to a certain income level; after this income level is reached, the positive relationship between income and happiness either drops or stays the same. However, the distribution of our data is heavily anchored on the poor; there is not enough data from the World Values Survey to show what happens with the relationship of income and happiness among individuals belonging to the extreme side of the non-poor region.

5. ACKNOWLEDGMENTS

We thank Marites M. Tiongco, Pedro A. Alviola IV, and Mitzie Irene P. Conchada for helpful comments at various informal sessions.

6. REFERENCES

Easterlin, R. (1974). Does Economic Growth Improve the Human Lot: Some Empirical Evidence. New York Times. Retrieved February 26, 2014, from http://graphics8.nytimes.com/images/2008/04/16/business/Easterlin1974.pdf

Easterlin, R. A.; McVey, L. A.; Switek, M.; Sawangfa, O.; Zweig, J. S. (2010). "The happiness-income paradox revisited". Proceedings of the National Academy of Sciences 107 (52): 22463. doi:10.1073/pnas.1015962107

Guardabascio, B., & Ventura, M. (n.d.). Estimating the dose–response function through a generalized linear model approach. Retrieved October 29, 2014, from http://intra.itm.kth.se/polopoly_fs/1.472018!/Estimating the dose–response function through a generalized linear model approach..pdf

Kahneman, D. & Deaton, A. (2010). “High Incomes Improves Evaluation of Life but Not Well-being.” Proceedings of the National Academy of Sciences 16489–16493, doi: 10.1073/pnas.1011492107

Largoza, G.L. (2014). “The Easterlin Paradox Re-examined Using Propensity Score Matching on Philippine Social Weather Station Data.” Proceedings of the 2014 DLSU Research Congress.

Lopez, T. (2012, December 27). The Philippines is the happiest country on Earth. The Manila Times.

Rubin, Donald (2005). "Causal Inference Using Potential Outcomes". J. Amer. Statist. Assoc. 100 (469): 322–331.

Rubin, Donald (1974) "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies", Journal of Educational Psychology, 66 (5), pp. 688–701.

Stevenson, B. & Wolfers, J. (2008). "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox". Brookings Papers on Economic Activity, 2008 (Spring): 1–87.

Tan, M. (2005, September 7). Happy Filipinos. Philippine Daily Inquirer.