HAPPINESS: A POLICY PERSPECTIVE

By

Alexo Esperato-Martínez

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
February, 2014

ABSTRACT

Objectives: this thesis examines the relationship between government policy and life satisfaction (happiness). This is accomplished through three related aims: a) to determine whether subjective or objective health are related to happiness for people above age 60; b) to ascertain whether health during childhood affects happiness after age 60, and c) to evaluate the impact of a senior citizen program (Golden Citizen) on happiness.

Methods: *e*conometric analysis of a nationally representative population over age 60 in Costa Rica (n=2,827). The availability of two datapoints permits the use of various econometric techniques. The first two aims rely on cross-sectional and lagged models (first differences, lagged dependent variable models and change-scores). The last aim relies mainly on quasi-experimental techniques, including instrumental variable, regression discontinuity, and differences-in-differences. Additional sensitivity analyses are provided for each of these aims.

Results: *a)* subjective health predominates over objective health in predicting happiness in people over 60; b)early life health is related with happiness after age 60, but only in its subjective variant; c) the Golden Citizen program has consistent effects on poverty, inconsistent effects on healthcare access, but no effect on happiness.

Conclusions: subjective health predominates over objective health in predicting happiness. This is true for both child and adult health measures. Policy-makers should focus on subjective health predictors (depression, pain and others) if health policy is going to truly impact happiness. Given its importance to most individuals, happiness should be

an explicit outcome in impact evaluations. The development of a valid happiness measure is crucial to future research.

Thesis readers: Gerard Anderson, Donald Steinwachs, Elizabeth Stuart, and Mariana Lazo.

PREFACE

Acknowledgement

As any research endeavor, this doctoral thesis has been the effort of more than one person. The author would like to specially thank Dr. Gerard Anderson for his invaluable role as a doctoral adviser, and the thesis committee (Dr. Elizabeth Stuart, Dr. Donald Steinwachs, Dr. Antonio Trujillo and Dr. Mariana Lazo for their comments on various drafts. Finally, the author would like to thank the department of Health Policy and Management and the Bloomberg School of Public Health for their generous support.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Everyone would like to be happy. In a survey of 7000 people across 42 countries, 69% of respondents placed happiness as their top priority [1]. Such prominence has compelled academics to consider happiness as the "ultimate aim of human endeavor"[2]. In addition to having intrinsic value, mounting evidence supports that happiness can improve various life domains. Happier people heal faster[3], are less likely to get infected, more resistant to common diseases[4], and enjoy better overall health[5]. Learned optimism, a closely related concept, is protective mental disorders like depression[6]. Likewise, there are indications that happy individuals have a greater ability to earn money [5]. Therefore, the pursuit of happiness is both an end, and a means to improve other life dimensions.

Politicians are also concerned with happiness. For US President John Adams, "the happiness of society is the end of all government"[7]. Under the auspices of David Cameron, the prime minister of the United Kingdom, the Office of National Statistics of the UK has just published its first national well-being statistics [8]. In France, former Prime Minister Nicolas Sarkozy convened a commission of experts to identify the limits of "GDP as an indicator of economic performance and social progress". This commission recommended shifting policy emphasis from economic production to wellbeing[9], a concept akin (and used often interchangeably) with happiness. Along the same lines Germany, Australia and China have elevated "happiness to the status of official government policy, drawing on the academic happiness revolution to define their goals and metrics"[10]. Inter-governmental organizations are following suit. The European Union has announced its intention to complement its GDP statistics with wellbeing figures, and the Organization for Economic Cooperation and Development (a think tank of 33 wealthy countries) just launched its "Better Life Initiative", which will compare subjective wellbeing across countries[10].

The academic community has followed suit. To be sure, the happiness phenomenon has been discussed since antiquity. Aristotle, one of the most influential thinkers in Western civilization, considered happiness to be the "the only good that is good in itself"[11]. Since then, a wealth of thinkers have occupied themselves with the happiness construct, including Seneca, John Stuart Mills, Jeremy Bentham[12], and Bertrand Russell[13]. However, the data-driven study of happiness did not take off until the foundation of positive psychology at the end of the 1990s [14]. The extent of this movement—which spans psychologists, economists, epidemiologists and political scientists—has compelled some authors to speak of an academic "happiness revolution"[10]. Indeed, the first World Happiness Report[11], published last year, is co-authored by Jeffrey Sachs, and includes contributions of Nobel awardees such as Daniel Kahneman. In the same vein, the National Science Foundation in the US has recently funded a project to develop a new happiness index[10].

Until recently, it was widely believed that economic development brings happiness automatically. This assumption is rooted in classical economic theory, which underscores capitalism's capability to improve utility, a related concept. Economic development brings about food, shelter, and security; the transition from a "society of scarcity to a society of security" generates abundant wellbeing.[15] Indeed, the relationship between income and happiness is strongest in low-income countries[16]—where scarcity is more common. Further theoretical frameworks postulate that economic development can also increase happiness also in developed countries. The most salient of these is human development theory, which postulates that as countries transition to higher levels of abundance, income affects happiness through free choice[15]. Indeed, a universal link exists between happiness and free choice[17]; by fostering a wider range of choices, economic development could contribute to happiness also in rich countries[15]. Taken together, these theories suggest that the main governmental instrument to foster happiness is economic growth.

However, such assumption is under review. First, setpoint theory—which is well-established in psychology—postulates that individuals have a fixed happiness level to which they revert inevitably. Life events such as changes in income may initially affect happiness, but individuals will eventually return to their personal baseline (Headey & Wearing, 1992). Indeed, people adapt to income[18],[19] through a “hedonic treadmill”: as income rises, aspirations rise accordingly [20,21], so in the long run income will not necessarily bring happiness[22],[23]. The partial disconnect between income and happiness is further reinforced by the Easterlin Paradox, which states that although richer people tend to be happier, average happiness does not increase with rising average incomes[24]. Indeed, this paradox has been confirmed in the US[25] and 35 other countries[24]. The most accepted explanations for the Easterlin paradox are reverse causality (i.e. happier people are more able to make money) and social comparison (richer people are happier because they compare themselves to those worse off) [5,24,26]. Others argue that income does add to happiness, but with diminishing returns. Indeed, some have found a logarithmic[11,11,27]relationship between income and happiness: after a certain threshold is reached, income may contribute little to happiness; this threshold has been set at 70,000 dollars in the US[28].

Economic growth also levies substantial costs. These, discussed in the World Happiness Report[11], can be grouped into three categories: individual infirmity, deterioration of community relationships, and deterioration of the environment. The first, also labeled as "the ills of modern life", comprehend man-made conditions and psychosocial disorders. Man-made conditions include obesity and its related epidemics (type two diabetes and others), which started in Western countries, and are now being exported to the developing world[29]. The psychosocial disorders include depression, anxiety, and eating disorders, all linked with stress[30], a landmark of high growth societies. The deterioration of social relations seems to occur through increased competitiveness, reduced dependency, and isolation. Competitiveness can weaken social trust; wealthy, highly technological societies allow individuals to fulfill many basic needs (clothing, nourishment, sheltering, etc.) without depending on their immediate community. The Internet and social media make it possible for individuals to socialize without leaving their households, which often results in less meaningful relationships. Finally, sustained growth demands sustained resources, which has resulted in an unprecedented scale of environmental destruction. It has been calculated that 1.52 planets are needed to sustain today's humanity; this is 2.61 times the resources needed in 1961[31]. The scientific community has long warned about this environmental depletion, which has been linked with the risk of natural disasters[32],[11], and the loss of 30 percent of the animal species[33].

In sum: income increases happiness among the poor, but once a certain threshold is reached, the connection between the two becomes tenuous, and rising incomes come at substantial costs. In the World Happiness Report, Oliver Sachs and others summarize these issues as the follows[27]:

"Now we face a set of real choices. Should the world pursue GNP to the point of environmental ruin, even when incremental gains in GNP are not increasing much (or at all) the happiness of affluent societies?
Should we crave higher personal incomes at the cost of community and social trust? Should our governments spend even a tiny fraction of the $500 billion or so spent on advertising each year to help individuals and families to understand better their own motivations, wants, and needs as consumers?[11]"

The study of happiness considers how the whole human experience affects wellbeing. It considers many policy-relevant dimensions beyond income, such as health, social relationships, and civil status. Given its broad scope, embedding such findings in policy-making can bring substantial benefits. First, policies would be explicitly aligned with their intended bottom line (i.e. happiness) of most individuals. Second, countries could strike a wise balance between economic growth and other policies. For example, accumulating research suggests that non-pecuniary life dimensions (marriage, health, retirement) may exert durable effects on happiness[34]. Therefore, each country would need to find the best combination of GDP growth and these other dimensions. Third, happiness can provide a common metric to compare policies. In the current situation, policies in different sectors (i.e. health, transportation, education) cannot be directly compared, because their outcomes are sector-specific. If evaluated in terms of happiness, these policies can be compared head-to-head, allowing policy-makers to find their best combinations, and to integrate them around their common goal. Indeed, prominent researchers are advocating for the use of happiness as an evaluation criterion[11]. Fourth, promoting happiness amounts to promoting health. Wellbeing is a necessary component of health, which is commonly defined as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"[35]. This comes in addition to the health benefits discussed at the beginning of this chapter, such as happier people healing faster, and happiness as a protective factor from mental disorders. The fifth benefit of happiness has deeper philosophical roots, and concerns life satisfaction, which is the measure used in this doctoral thesis. Traditionally, philosophers and psychologists have contrasted the hedonistic (i.e. pursuit of pleasure) with the eudemonistic (i.e. pursue of excellence) lifestyle, where the first emphasizes "positive emotions" and the eudemonistic a "sense of purpose and meaning". Life satisfaction questions have the virtue of taking both into account, providing clues about the deeper aspects of a good life[11]. If embedded in policy-making, happiness research can thus bring substantial benefits to the whole human experience, and at lower personal, community, and environmental costs.

Given the benefits of happiness, should governments pursue the happiness of their citizens? Two schools emerge on this point. The social-democratic school assumes that governments can increase happiness by maximizing societal welfare[7]. In this view, government action would be particularly beneficial to vulnerable groups, such as the elderly, which is the target population of this study. In contrast, the Anglo-Saxon liberal view regards governments as "inefficient and self-serving" [7] that inherently alienate their citizens. This happens through "collectivization": government action displaces important societal structures such as church and family, and curtails the autonomy of individuals to pursue their own happiness. In consequence, happiness will drop. Therefore, the study of happiness is also highly relevant to the role of the state.

Happiness research has shed new light on the relationship between happiness and policy—and hence happiness and the state. This field argues for a paradigm shift, which, fueled by the availability of new data, challenges two traditional assumptions in psychology and economics. First, positive psychology—the precursor of happiness research—emphasizes a paradigm shift from a "preoccupation only with repairing the worst things in life to also building the best qualities in life"[6]. Indeed, the focus on pathology "neglected the idea of a fulfilled individual and a thriving community, and it neglected the possibility that building strength is the most potent weapon in the arsenal of therapy" [6]. Second, a growing number of economists are questioning the revealed preferences approach to understanding utility. This approach places heavy emphasis on the biases (cognitive and others) of self-reported outcomes, such as happiness. Therefore, and under the assumption that individuals choose the combinations of goods and services that maximizes their utility (happiness,) economists have traditionally argued that the only way to know what makes people happier is by observing their actual choices. However, new research suggests that individuals are not always able to choose what makes them happier, due to the influence of context[36], ethical convictions[36], or their powerlessness to influence specific policies or institutional arrangements[12]. Hence, economists are increasingly using self-reported outcomes, such as happiness.

Though still in its infancy, happiness research is growing quickly, and can point to a few points of consensus. First, genetics explains a large share of happiness; studies on twins show it to explain between 30%-80% of individual happiness [37]. Second, other important happiness predictors include health, income, civil status, employment, and social ties [38-40]. Amongst these, health seems key [41,42,43,44,], particularly in the elderly[45]. Third, new evidence is increasingly challenging the setpoint theory. A substantial body of research supports that individuals do not always adapt to their baseline happiness after major life events such as divorce[46], widowhood[47], [48], disability[49], and unemployment[50]. At the same time, the happiness literature exhibits some gaps. First, it has mainly relied on data from developed countries. Developing countries are different in many dimensions, such as culture (developed countries, with the exception of Japan, belong to Western cultures) and economic stability. Hence, the findings of happiness research from the rich countries may not always translate to the developing world. Second, happiness studies that use longitudinal nationally representative databases are still scarce in developing countries. The use of such data would allow adjusting for the effect of unobservable factors, and extrapolating findings to a national scale. Third, happiness in the elderly has been seldom studied, and when so, existing research tends to on small-scale surveys[51-55]. Given their vulnerability status and the ageing of developing countries, more research on the elderly is advisable. Fourth, as in any field, important questions are still in need of scrutiny. These, all in the health sector, can be summarized as follows:

- Health is important predictor of happiness; however it is unclear whether it is objective or subjective health that predicts happiness. In the first case, traditional health policy, which focuses on objective health, would offer a direct venue to maximizing happiness. If it is subjective health that affects happiness, health interventions may need redesign to truly affect happiness.
- There is still limited understanding on the relationship between early life health and senior happiness. Lifecourse studies have documented the effect of various adverse childhood circumstances (i.e. health[56], poverty[57], depression) on adult health. Similarly, it can be postulated that adversity at young age could affect long-term happiness. To our knowledge, these relationships have been explored in only one paper[58].
- Program evaluations in terms of happiness are nonexistent in developing countries.

This doctoral thesis addresses the previous three questions. Chapter two explores which dimensions of health (subjective and objective) predict happiness. Chapter three examines the relationships between childhood adversity (health and poverty) and happiness at late age. The fourth chapter evaluates the impact of a health and social services program (the Golden Citizen program) on happiness. Finally, a concluding chapter summarizes lessons learned and offers implications for future research. Taken as a whole, these three articles examine whether (health) policy can affect happiness, and hence whether governments can promote happiness.

All the analyses use a nationally representative survey of an elderly population in Costa Rica, a middle income country. The longitudinal data structure permits overcoming previous methodological difficulties, such as adjusting for the effect of unobservable factors, statistical precision concerns, and extrapolating the findings to the national scale. The focus on the elderly seems justified: Latin America—and particularly Costa Rica--is projected to age rapidly over the next decades[59].

By exploring the relationship between health and happiness, this thesis will hopefully contribute to policies that improve people's health, are more attuned with people's ultimate goals, and incur lower costs. Understanding these questions will serve a better understanding of the ultimate question–whether the state can promote happiness.

2. THE RELATIONSHIP BETWEEN HEALTH AND HAPPINESS

Abstract

Health is a major predictor of happiness. However, it remains unresolved whether subjective or objective health accounts for this association. Using a nationally representative dataset in Costa Rica (CRELES), this study finds that subjective health predominates over objective health in predicting life satisfaction (a measure of happiness) in people over 60. While, adjusting for confounders, all subjective health measures (depression, pain and perceived health in two different variants) are correlated with life satisfaction ($p<0.05$), the only significant objective health dimension was disability ($p<0.10$). The subjective measures are two to four times more important predictors of life satisfaction than disability. Our findings suggest that if policy makers want to use health care to improve happiness, they should concentrate on subjective measures, particularly depression and pain.

Introduction

Happiness has been always a human pursuit. Its predominance in human thinking has compelled some academics to name it as the ultimate human endeavor [2]. In a major international survey, 68% of respondents rated happiness as their highest priority [60]. The pursuit of happiness has occupied philosophers and economists for centuries, including Aristotle, Bentham, John Stuart Mills and Adam Smith [12] (Graham, 2005). The United States Declaration of Independence published in 1776 considers the pursuit of happiness to be an inalienable right. More recently the Prime Ministers of the United Kingdom and France are exploring how their governments can increase the level of happiness, rather than just the economic well-being, of their citizens [9,61]. In Bhutan, the constitution urges the State to "strive to promote those conditions that will enable the pursuit of Gross National Happiness" [61]. Such developments suggest that the improvement of happiness is a plausible government objective.

The recent decade has witnessed an explosion of happiness research. Some points of consensus emerge from this literature. First, the relationship between age and happiness is U-shaped: the young and the elderly tend to be happier than the middle-aged. These findings are present in developed [62-64] and developing countries [62,65,66]. Second, four key dimensions, which change along the life course, seem particularly associated with happiness: income [38,67-69], civil status [43,45,70], employment[38,71,72], and health [41,43].

Amongst the previous factors, health may be the most correlated with happiness. However, "it is less clear whether objective or subjective health is responsible for this association" [39]. Measuring and improving objective health is where most of the policy and clinical attention has focused. For example, considerable policy and clinical attention has been given to whether diabetics receive appropriate care or how hospital safety can be improved. If objective health matters, then health care offers a direct venue to happiness.

If it is subjective health that affects happiness, then the relationship between health care and happiness is more indirect. This is because subjective health is strongly affected by factors beyond health care, such as psychological traits and socioeconomic status [73]. In such case, interventions such as improving care for diabetics or increasing hospital safety may not be enough to increase happiness.

To answer whether subjective or objective health predict happiness, the current study analyzes a nationally representative sample of senior citizens in Costa Rica, a middle income country with growing population +60 years old [74]. As the determinants of happiness tend to be similar across countries, this study has implications for high and other middle income countries.

Methods

Conceptual framework

This paper explores whether subjective or objective health predict happiness. To do so, we developed a general happiness model that regresses happiness (life satisfaction) on health (objective and subjective) and a set of covariates selected from the current literature. These include the main happiness correlates (income; health; civil status); a set of exogenous variables that are usually related to happiness (age, sex, educational level, number of people in the household, urban residence)[38], personality (locus of control) [75,76], cognitive skills [77], exercise, and two variables of particular relevance to Latin America - how often the respondent sees his or her children [78], and weekly church attendance [79].

Health and happiness are also affected by time-invariant unobservable characteristics, such as genetics, personality, culture, and life events. These unobserved factors are crucial: genetics alone explains between 33%

and 80% of happiness variability across people [37]. Hence, controlling for such factors is crucial in happiness research [80]. We attempt to control for these through econometric analysis.

Data

The data source for this study is the Costa Rican Study on Longevity and Health Aging (CRELES). This is a nationally representative panel dataset of individuals 60 years and older in Costa Rica. Its main objective is to determine "the length and quality of life, and its contributing factors in the elderly"[81]. CRELES collected data on self-reported physical health, psychological health, living conditions, health behavior, health care utilization, social support, and socioeconomic status. The survey conducted two rounds of interviews. The first, completed between November 2004 and September 2006, includes 2,827 people. The second, completed between November 2006 and July 2008, contains 2,364 individuals from the first round. Attrition over the two rounds was therefore 16% (463 people). The majority of the censored individuals (10%) were deceased by the second interview, and the remainder (6%) could not be located [82]. Bivariate analyses revealed that censored and non-censored individuals were not statistically different in happiness, socio-demographic, and health variables. Hence, attrition-generated bias should be modest.

Modeled after the Health and Retirement Survey from the United States, CRELES uses a complex design. A master file of 9,600 individuals born before 1946 was first constructed. This was accomplished by stratifying all 55+ individuals from the 2000 Census of Population into 5-year age groups; random sampling from these strata, and oversampling the oldest-old (95 years and older). Sampling fractions varied between 1% (for the ones born between 1941 and 1945) and 100% (for the born before 1905). The 9,600 individuals in the master file were then clustered into the government's already designated 102 Health Areas. A probabilistic sub-sample of 60 of these clusters was selected for the first interview. The resulting sampling frame contained approximately 5,000 individuals, covering about 59% of the national territory. Of these, 2,827 were located and interviewed in the first wave. Non-responses (43% of sampling frame) occurred for the following reasons: 19% of potential respondents were dead, 18% could not be located (mainly due to lack of accuracy in existing addresses), 2% had moved, and 4% declined (directly or indirectly) to be interviewed [81]. The majority of non-respondents were younger, of different social class, and lived in different cities than the remainder of the sample. In order to correct for oversampling and non-response [83], CRELES incorporates a set of weights. These "allow the replication of the structure for sex, age, residence and education of the whole 2005 population of Costa Rica born in 1945 or before" [81]. Our statistical analyses take these weights into account.

To validate the data, the survey team compared CRELES to the Costa Rican Household Survey for Multiple Purposes (EPHM) on nine key measures (age, education, index of masculinity, % head of household, and others). EHPM is a nationally representative survey of 12,000 households conducted yearly by the National Statistics and Census Institute of Costa Rica [81]. The two surveys differed in only one variable (% head of household). This difference, which is significant at the $p<0.05$ level, was attributed to the fact that in "CRELES the informant is the same older person while in EHPM it can be another person with a different perception on who is the head of the household." [81]. The coherence between both surveys suggests an “absence of biases that could have damaged the representativeness of the sample” [81].

Another potential source of bias, typical of elderly populations [84], lies in the use of proxy interviews. At CRELES, 703 interviews in the first wave, and 676 in the second necessitated a proxy due to incapacity to communicate or cognitive impairment of the main respondent [73]. Such cases contain no self-reported variables, including happiness. These observations were excluded from the analytic sample. Given that cognitive impairment was the main reason for proxy interviews, and that answering life satisfaction questions requires interviewee awareness of their current situation, the exclusion of these individuals from our analysis is necessary. Therefore, the final study populations includes 2,111 observations in the first round; 1,684 observations in round 2.

Variables

The dependent variable is life satisfaction. Life satisfaction is one definition of happiness, along with positive affect, subjective well-being, and other measures [85]. A key distinction between these is stability[39]. Life satisfaction questions require a cognitive evaluation of one’s situation, it is therefore more stable than the other happiness measurements. Thus, life satisfaction is a standard happiness measure in large-scale studies [44,49,58,72,86-88]. Our dependent variable was extracted from the most commonly used happiness question [12,89]: “in general, how do you feel about your life?: very satisfied, somewhat satisfied, somewhat unsatisfied, very unsatisfied”. This question was asked in both waves. In a manner similar to other countries, responses to this question presented a ceiling effect—74.7% of respondents in wave 1, and 77.8% in wave 2 reported feeling very happy. Power calculations determined that variability in some categories of the outcome would not be sufficient for statistical inference. As a result, the dependent variable was dichotomized into being very happy or less than very happy.

There are two categories of health measures: subjective health and objective health. Subjective health can adopt multiple dimensions [90]: the respondent's description of overall health (i.e. self-reported health status; pain), mental feelings (depression), and social feelings (i.e. feels valued by others). At CRELES, respondents reported their self-perceived health in four different questions. Two of these are available in both waves: self-reported health ("How would you say your health is now: Excellent, Very, Good, Good, Fair, Poor") and comparison of one's health to his/her peers ("How would you say your health is in comparison with other people of your age? Better, Equal, Worse?). Both variables were dichotomized along the good versus fair/poor axis. Respondents were also asked if they had suffered daily pain over the last year in the stomach and in the lower limbs (yes/no). The depression variable (dichotomous) reflects if an individual is clinically depressed according to the 15-item Yesavage Geriatric Depression Scale (GDS-15). Following a psychometric validation, an individual was classified as depressed if they had ten or more, and as mildly depressed if they had five or more depression symptoms[91].

Objective health measures are "concerned with form and function either of the whole body or of its constituent organs" [90]. Hence, they refer to externally measured physical health. In our analyses, objective health measures are composite variables, derived from individual measurements taken during the interviews. The objective health measures are: disability, chronic diseases, allostatic index, and metabolic syndrome.

The disability and metabolic syndrome variables were coded by the CRELES team [92]. Disability was recoded from 14 items on the Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL). The 10 ADL questions assessed difficulties with basic functioning (walking, bathing, eating, going to bed, using toilet, and cutting toenails), and the four IADL questions measured any difficulties with more complicated but accessory tasks (preparing food, managing money, shopping, and taking medications). As a sum of these items provides a reliable measure [93], we added them and a person was coded as disabled if they could not perform 5 or more of these activities. This cutoff had been proposed by the survey team[92]. The metabolic syndrome variable is based on the criteria of the International Diabetes Federation [94]. Thus, individuals were classified with metabolic syndrome if they had abdominal obesity (waist-hip ratio>1) and any two of the following: high triglycerides (>150 mg/dL); low HDL cholesterol (<40 mg/dL in males; < 50 mg/dL in females); hypertension (>130 systolic; > 85 diastolic), or diabetes. The chronic disease variable denotes having at least three chronic diseases as measured by a count of self-reported questions on cancer, heart attack, heart disease, stroke, hypertension, high blood lipids, diabetes, lung disease, arthritis, and osteoporosis. Allostatic load represents the 'wear and tear' of the body resulting from the daily stress [95]. Hence, the higher the allostatic load, the higher the wear and tear, and the worse the health. The allostatic load measure is a sum of 7 dimensions: systolic blood pressure, diastolic blood pressure, glucose levels (hba1c), waist-hip ratio, high density lipoprotein cholesterol (HDL) and Dehydroepiandrosterone (DHEAS). An individual gets a point for each dimension if he/she was in the top 25 percentile of the distribution [96]. Validation studies [95] found allostatic index to be predictive of cardiovascular disease, cognitive impairment, mortality, and physical decline [95,97].

The analyses use a set of additional covariates from the happiness literature. These include age (5-year age groups); sex; civil status (=1 if married or cohabiting), educational level (4 categories); number of household residents (continuous); urban residence (dichotomous); health region code (7 health regions); weekly church attendance; weekly visits with children; weekly exercising; cognitive impairment (=1 if suffering cognitive impairment); locus of control; living below the poverty line; and housing assets. Cognitive impairment (continuous, range 0-100) was measured through the PFEFFER scale[81]. Locus of control reflects the degree to which the respondent's feels in control of their lives. The locus of control score sums 7-items; higher scores reflect higher control [98]. Determining poverty levels required first developing income variables. CRELES contains questions on income of the respondent and their spouse. In both cases, personal income was calculated as the sum of income from work, pension, and transfers from others. Given the absence of data for other household members, total household income was the sum of personal respondent and the spouse's income. Afterwards, household income was "equivalised", i.e. adjusted to family size. Because information was only available for the respondent and their spouse, household income was "equivalised" for two people, by dividing it by $\sqrt{2}$ if the respondent was in a relationship [99]. Such adjustment reflects that each of two people in the same household consumes less than a person living alone. Locative values (i.e. how much the respondent would receive from renting the house they own and live in) were then added to "equivalised" incomes. This adjustment is standard in large national surveys [100]. After these adjustments, poverty rates were calculated by applying national 2004 and

2007 poverty thresholds [101] to the calculated incomes. Finally, an assets index provides a measure of permanent income [102]. This index is a weighted sum [103] (range 0–32) of a set of questions on housing quality (quality and materials of floor, ceiling) and home assets (room for cooking, number of telephones, of computers, etc.). In this approach, each asset is assigned a weight, which is inversely related to the number of respondents owning that asset. Hence, asset index=∑asset*w, and a higher score reflects higher assets.

Imputation of missing data

At CRELES, missing data is not a significant problem: only 5% of observations were missing two or more happiness predictors. Variables with missing data were imputed. Only three covariates (depression scale, locus of control, and allostatic index) were missing in >5% of observations. Imputation proceeded through two different methods. First, single imputation proceeded for index variables (i.e. variables that are the sum of subvariables). These are depression, disability, locus of control and assets. Single imputation was appropriate as information was available on the majority of subvariables in each index (only <5% of observations were missing >20% of items in the index). The rest of variables (chronic diseases, frequency of seeing children, metabolic syndrome) were imputed through multiple imputation chained equations (MICE)[104]. MICE predicts missing values based on other variables with non-missing values; it imputes first the variable with the least number of missing observations; then the second, and so on. All variables are cycled through in such manner, and the whole process is repeated (usually 5-10 times)until a dataset with converged estimates has been reached. This process produces a single imputed dataset with no missing data[105].

Statistical methods

Exploratory data analyses examined the analytic sample characteristics. Then, the relationships between happiness (dependent) and subjective and objective health (main independent variables) were explored. The objective here was to examine if, before adjusting for confounders, the main independent variables were associated with happiness.

Because a large portion of happiness is due to genetics, first-differences (ΔY on ΔX) models are the main analysis for this study. As a sensitivity analyses, we also conducted a lagged-dependent variable model or LDVM (Y_2 on X_2 and Y_1). We also report change-scores (Y_2-Y_1 on X1) and pooled cross-sectional as additional sensitivity analyses. As explained in the following sections, the first three approaches cancel out (at least partially) the influence of time-invariant unobservable characteristics, including genetics.

First-differencing

First differences (FD) models regress changes in the outcome on changes in the predictor variables. Hence, they adopt the following general form:

$$\text{probit}(\Delta \text{Life Satisfaction})=\beta_0+ \beta_1(\Delta \text{ Self-Reported health})+ + \beta_2(\Delta \text{Disability})+ \beta_3(\Delta \text{Chronics})+ \beta_4(\Delta \text{Allostatic Load})+ \beta_5(\Delta \text{Metabolic Syndrome}) +\beta_6(\Delta \text{Depression})+ \beta_7 (\Delta \text{Cognitive Impairment}+ \beta_8 (\Delta \text{Age})+\beta_9 (\Delta \text{Civil status})+ \beta_{10}(\Delta \text{Household residents}) +\beta_{11}(\Delta \text{Attends Church})+ \beta_{12}(\Delta \text{Sees Children})+ \beta_{13}(\Delta \text{Exercise}) +\beta_{14}(\Delta \text{Poor})+ \beta_{15}(\Delta \text{Assets}) + a_i+u_i$$

(equation 1)

In the previous equation, the error term is decomposed into time-invariant (a_i) and time-variant unobservable factors (u_i). First-differencing cancels out the effect of time-invariant factors (a_i)[106]. These include genetics and personality, which are crucial in happiness research [80]. Following the approach from Roberto and Gaskin [107], changes in life satisfaction were modeled in the following form: 0 for persistent dissatisfaction (less than very satisfied at both time points), 1 if satisfaction decreased between both waves, 2 if satisfaction increased, and 3 for persistent satisfaction(i.e. individuals very satisfied at both waves). These are ordered from the least to the most desirable outcome. The categorical predictors were modeled in the same manner (4 point scale). These estimations used an ordered probit model with partial proportional odds; i.e. proportional odds for individual variables are assumed only if supported by the data. This was attained through gologit2-, a specialized command in Stata [108] that tests and empirically applies the proportional odds assumption. The partial proportional odds approach enhances the parsimony, and therefore the reliability of estimates. The first-differences analyses express results in two ways. The first is predicted probabilities, that is, the probability change in a given level of the outcome resulting from a one unit-increase in the independent variable. The second is standardized coefficients. These reflect the standard-deviation change in the outcome resulting from a one-unit standard deviation change in the independent variable. While predicted probabilities provide an intuitive interpretation, standardized coefficients provide a common basis to compare estimates. Given their easier interpretation, the results are mainly interpreted in terms of predicted probabilities.

In order to explore whether ADLs are indeed more strongly associated with happiness than IADLs, we ran two additional first difference specifications. The first follows the equation above, and conceives disability as transitioning into or out of suffering at least one ADL (4 categories). The second does the same for IADLs.

Main sensitivity analysis: lagged dependent variable model (LDVM)

Econometric researchers are concerned about bias in first-differences models that use ordered probit [80,109]. Therefore, a lagged dependent variable model (LDVM) is used as the main sensitivity analysis. This specification regresses second round happiness on round two covariates and first round life satisfaction (i.e. Y_2 on X_2 and Y_1). Accordingly, it takes the following form:

probit(Life Satisfaction$_2$)=β_0+ β_1(Self-Reported Health$_2$)+ β_2(Disability$_2$)+ β_3(Chronics$_2$)+ β_4(Allostatic Load$_2$)+ β_5(Metabolic Syndrome$_2$) + β_6(Depression$_2$)+ β_7(Cognitive Impairment$_2$)+ β_8 (Age$_2$)+ β_9(Sex)+ β_{10}(Civil status$_2$)+ β_{11}(Educational Level)+B$_{12}$ (Household residents$_2$) +β_{13}(Area of residence$_2$) + β_{14}(Attends Church$_2$)+ β_{15}(Sees Children$_2$)+ β_{16}(Exercise$_2$) + β_{17}(Poor$_2$)+ β_{18}(Assets$_2$)+ β_{19}(Life Satisfaction$_1$)+ε (equation 2)

By including baseline life satisfaction, LDVM limits reverse causality and some unobserved confounding [110]. Typically, LVDM adjusts for some “historical confounders”, i.e. variables associated with both health and happiness [106]. However, LDVM models are not as efficient in eliminating time-invariant unobservable factors (i.e. genetics, personality) as first difference models.

In this study, the estimations that rely on a binary outcome (LDVM and pooled cross-sections; see next section) report two kinds of coefficients:

marginal effects and standardized coefficients. Marginal effects express the average probability change in the outcome that results from a one unit increase in the independent variable. Standardized coefficients, as explained earlier, report the standard-deviation change in the outcome resulting from a one-unit standard deviation change in the independent variable. Given their easier interpretability, this article focuses on marginal effects; however the reader is encouraged to check the standardized coefficients to compare the estimates.

Additional sensitivity analyses: pooled cross-sectional and change-scores

We also report pooled cross-sectional and change-scores. The pooled cross-sectional model uses the same covariates as in equation 2 and controls for time effects through a year dummy. Pooling both waves increases the sample size and thus statistical precision, versus the alternative of running separate cross-sections. However, as any cross-sectional model, pooled models are vulnerable to omitted variable bias [111]. A separate pooled model proceeded for each perceived health measure available (self-reported health; health compared to peers). As, the pooled cross sections use also a binary outcome, coefficients are reported in both marginal and standardized forms[1].

The change-scores models regress the changes in life satisfaction on the baseline predictors (ΔY on X_1):

$$probit(Life\ Satisfaction_{2-1})=\beta_0+\beta_1(Self\text{-}Reported\ Health_1)+\beta_2(Disability_1)+\beta_3(Chronics_1)+\beta_4(Allostatic\ Load_1)+\beta_5(Metabolic\ Syndrome_1)+\beta_6(Depression_1)+\beta_7(Cognitive\ Impairment_1)+\beta_8(Pain\ Stomach)+\beta_9(Pain\ Lower\ Limbs)+\beta_{10}(Age_1)+\beta_{11}(Sex)+\beta_{12}(Civil\ status_1)+\beta_{13}\ Educational\ Level)+B_{14}(Household\ residents_1)$$

[1] FD and change-score models were estimated through a specialized Stata command (-gologit2-). This command does not allow the estimation of standardized coefficients. Therefore, standardized coefficients are only reported for LDVM and pooled models.

$+\beta_{15}(Area\ of\ residence_1) + \beta_{16}(Attends\ Church_1) + \beta_{17}(Sees\ Children_1) +$
$\beta_{18}(Exercise_1) + \beta_{19}(Poor_1) + \beta_{20} + (Assets_1) + \varepsilon$ (equation 3)

These models use an ordered probit specification and also conceive changes in the outcome as a 4-level category. Conceptually, they are an alternative to LDVM; however the econometric literature considers them to be more prone to estimation problems [110]. As (unlike first-differences and LDVM) change-scores rely on baseline predictors, they permit estimating the relationship between happiness and those subjective health measures available in round one only, which include pain (stomach/lower limbs). Therefore, change-scores is the sole method to estimate the relationship between happiness and pain.

All the previous models use self-reported health as the perceived health variable. An additional appendix provides estimations for the other perceived health measure in our survey: health compared to peers. As this variable is also available in both waves; it is also assessed through FD and LDVM specifications.

Mediation analyses for policy groups

Mediation tests confirm whether perceived health could mediate between happiness, on the one hand, and depression/poverty on the other. This is warranted because perceived health has been shown to be affected by the last two dimensions [73]. If our data support perceived health's mediating role, then interventions on poverty and depression could improve perceived health (i.e. mediator) and hence happiness. The mediation tests were run on the LDVM models in equation 2. LDVM was chosen to test mediation because of methodological limitations: standard commands in statistical packages cannot estimate mediation in variables with more than two levels; mediation tests in the first-differences model was therefore not possible, but they are in LDVM – the second most robust analytical method. In order to estimate the models, we used the binary mediation command in Stata 11 [112].

Most standard mediation tests—such as the one used in this essay—assume that the mediator (i.e. perceived health) is randomly assigned[113], and hence unaffected by unobservable confounders. However, the subjective nature of perceived health makes it likely to be affected by unmeasured variables (i.e. pessimism). Further, perceived health may affect the exposure variable (i.e. depression), which would also violate the previous assumption. The test assumes further that the main exposure (i.e. depression) affects the mediator (i.e. perceived health), which in turn affects the outcome (life satisfaction). However, causality could also run in the reverse direction: life satisfaction could affect self-perceived health, which could in turn affect depression. Hence, the results from our mediation test cannot be taken as conclusive. In all models, robust standard errors address heteroskedasticity concerns. Collinearity was assessed through variance inflation factors.

Results

As table 1 reflects, the characteristics of our sample did not vary substantially between both waves. The table shows two major exceptions. First, the percentage of individuals feeling very satisfied increased from wave 1 to wave 2 ($p<0.05$). Second, average incomes increased from 132,000 to 173,000 colones per month. The increase is consequence of general raises in pensions for the elderly introduced in 2006 and 2007[73]. Indeed, among those pensioned (54.4%) in our analytic sample, the 25% percentile went from earning 35,000 colones in round 1, to 52,000 colones at wave 2. The changes in the upper end of the distribution were even more substantial: the 75%-ile increased from 125,000 to 175,000 per month. In contrast, the average incomes for those not pensioned (45.6% of the analytic sample) did not increase between both interviews. The table also shows that, while objective health conditions deteriorated, subjective health improved. Therefore, subjective and objective health measures are not highly correlated, as they changed in different directions between the two time periods. Finally, average incomes increased substantially ($p<0.05$). This reflects a general increase in pensions for the poor between both time periods [73].

Table 2 disaggregates the bivariate relationships between health and life satisfaction. The table confirms that in both interview rounds the more satisfied group enjoyed better health -- subjective and objective. The only exceptions are allostatic load and metabolic syndrome; while the very satisfied still enjoy better health, differences are not significant. Therefore, this evidence suggests clearly that health is a strong correlate of life satisfaction.

The multivariate analyses suggest that subjective health measures are strongly associated with life satisfaction. We begin by analyzing the first-differences (FD) model. The predicted probabilities on table 3a show that a one-unit increase in self-reported health is associated with an 8.7% higher probability of persistent satisfaction ($p<0.001$) and a 4.3% ($p<0.001$) lower probability of persistent dissatisfaction. Depression is an even stronger predictor: a one-unit increase in depression is associated with a 23.6% decreased probability of persistent happiness ($p<0.001$), and 11.6% higher probability of persistent unhappiness ($p<0.001$).

Self-reported health and depression dominate also in all sensitivity analyses. The LDVM models in table 4 support that good self-reported health is associated with a 16.9% ($p<0.001$) higher probability of feeling very satisfied. Depression exhibits the largest relationship with life satisfaction: being depressed lowers the probability of feeling very satisfied by over 50% ($p<0.001$).

The cross-sectional pooled sensitivity analyses in table 5 replicate these findings: again, self-rated health and depression are the strongest life satisfaction. As in the LDVM models, these patterns are sustained when coefficients are both marginal (i.e. perceived health and depression are associated with the largest percentage changes in happiness) and standardized (i.e. when standardized to a common basis, these dimensions are still the strongest correlates of happiness).

The change-scores sensitivity analyses in table 6a sustain the same patterns: depression and self-rated health still dominate in predicting life satisfaction. Further, and as discussed in the methods section, change-scores were also used to estimate the associations of the two pain variables (upper and lower pain). The same table shows that both pain measures are highly statistically significant in predicting life satisfaction, and stomach pain has a stronger association with life satisfaction than pain in the lower limbs. Yet, the magnitude of the associations of pain with life satisfaction are about half of those of perceived health, which are much lower than depression. Taken together, the evidence is compelling in that all subjective health measures (i.e. depression, self-rated health, and pain) have strong and significant associations with life satisfaction.

Of all objective health measures, disability is the only dimension statistically associated with life satisfaction. However, its associations are of lower magnitude and intermittently significant. In the FD model (table 3a), disability results in a 2.6% ($p<0.1$) lower chance of persistent satisfaction and a 1.3% ($p<0.1$) higher chance of persistent dissatisfaction. Disability is also associated with happiness in both pooled cross-sectional models (table 6), where disabled individuals have a 6%-8% lower chance of feeling very satisfied ($p<0.05$). In the rest of models (bottom FD, LDVM, and change-scores), disability is non-significant.

The previous analysis suggests strongly that subjective health has far stronger associations with happiness than objective health. Indeed, the associations of subjective health measures (self-rated health, depression, pain) with life satisfaction are always far higher and more statistically significant than those for disability—the only significant objective health dimension. These findings are true for all the models just discussed, which presented results as predicted probabilities and marginal coefficients, thus allowing an intuitive interpretation. The standardized coefficients, which provide a more direct comparison of estimates, reinforce the same point: again, subjective health measures (depression, perceived health, and pain) have larger magnitude associations with life satisfaction than disability, which is the only significant objective measure. Table 3b presents standardized coefficients for the first-differences model; table 6b displays standardized coefficients for the change-scores model. Such findings are further buttressed by the specifications in the appendix, which use as an alternative perceived health measure the respondents' assessment of their health versus their peers'. The appendix portrays the same picture: subjective health prevails over objective health in predicting life satisfaction. Therefore, the findings are consistent in all analyses used in this study.

We then examined which measures of disability are associated with life satisfaction. ADLs predominate: in tables 7 and 8, the estimates for ADL are larger than for IADL. In addition, ADL estimates are always significant, whereas IADL estimates are significant in only one case. These associations seem plausible, given that ADLs gauge more fundamental life functions than IADLs [114] and thus imply a higher dependency on others.

The mediation tests elucidate whether poverty and depression may affect life satisfaction through perceived health (table 9). Poverty is not associated with life satisfaction, whether directly or through subjective health. However, depression does exhibit a strong association with life satisfaction, both directly ($\beta=0.363$; $p<0.001$) and through self-reported health ($\beta=0.058$; $p<0.001$). Indeed, the tests support that self-reported health mediates 13.1% of the relationship between depression and life satisfaction. Depression emerges therefore as a major determinant of life satisfaction, both directly and through self-reported health. However, as mentioned earlier the mediation tests must be taken with great caution given the likely endogeneity of our mediator variables (i.e. depression and others).

Generally, the rest of covariates perform as expected (results not shown in tables). Variance inflation factors across regressions always are below two; multicollinearity was therefore not a concern.

Discussion

This study suggests that subjective health is a far better predictor of life satisfaction than objective health. In other words: subjective health is strongly associated with life satisfaction, even after adjusting for objective health and other factors. This is the case for all subjective health indicators: depression, perceived health (two different variants) and pain (two different variants). Disability is the only significant objective health measure; its association with life satisfaction is also of lower magnitude and significance than the subjective health indicators. Within types of disability, ADLs are more predictive of life satisfaction than IADLs. Depression emerges as a key predictor of life satisfaction, and previous studies found them to be related [115]. Finally, the influence of depression on life satisfaction may be partly mediated by perceived health (i.e. perceived health mediates about 13% of the relationship between depression and life satisfaction).

These findings stand in contrast to some existing literature, which finds that physical health—particularly physical functioning—is an important predictor of elderly happiness[116,117-119]. However, the findings are in consonance with another major study on health and happiness in Latin America, which found anxiety and pain to be more important than physical health[44]. Adaptation may provide the most viable explanation for the predominance of subjective health: people can adapt to objective conditions. Indeed, some studies have found that individuals can adapt even to severe disability[49] Survivor effects may also explain the dominant role of subjective health. These imply that those who survive beyond a certain age are different on certain key characteristics than those who do not. For example, survivors may be more optimistic, and hence report higher subjective health and happiness than the average population. All individuals in our analytic sample survived until age 60; if survivor effects are present, then our sample may be comprised of individuals who are happier and tend to report better health--regardless of their objective health status. In such case, subjective health would become dominant. Existing literature suggests that survivor effects are possible: happier individuals tend to live longer [120,121] and be more optimistic [122]. However, our data do not support a survivor effect: adjusting for age, sex, and objective health, those who survived between both waves did not report better health than non-survivors (results available upon request). An additional explanation for the predominance of subjective health lies in social comparison, whereby people assess their health based on implicit comparisons with others. By comparing themselves with those in a more disadvantaged position, individuals may be giving more weight to their subjective (rather than objective) health [39]. As table 1 shows, the significant increase in perceived health measures between both rounds of interviews (chi2; $p<0.001$)-- while objective health tended to deteriorate—lends some support to this hypothesis.

The dominant role of subjective health offers two important implications for health policy. First, traditional health care, unless accompanied by other interventions, may have a limited effect on happiness. This is because health interventions target objective outcomes (i.e. levels of cholesterol, blood pressure levels, and so on), whereas perceived health has a wider set of determinants, some of which lie beyond the direct control of health care. Several studies have demonstrated that, across the world, people commonly report a different health state than as reflected by their objective health conditions[123]. This is largely due to health expectations, the use of the health services, and the understanding of health questions [123,124]. Perceived health is also more prone to be affected by social dimensions, such as household income [73,125] and family dysfunction [125]. Second, the strong effects of depression and pain suggest that treating these may indeed improve happiness. Hence, mental health and pain management may be adequate venues to affect the life satisfaction of seniors.

The current analysis presents limitations. First, life satisfaction is assessed through a single survey item. Most scientific traditions assume single-item variables to have lower reliability and validity than multi-item scales. However this is not necessarily the case in happiness measures[39]: previous attempts to validate multi-item happiness scales (e.g. PGC Morale Scale, SWLS, and PWI) have been unsuccessful, as the correlates of the scale may be confounded with the scale itself. A second limitation lies in the categorical nature of the outcome. Social desirability biases [126] and constraining the responses to four categories may result in a ceiling effect. When interviewed by a stranger, respondents may not admit to feeling "somewhat satisfied" or lower; individuals may therefore feel compelled to report feeling very satisfied. Ceiling effects are typical across the world; in the US, 80-85% of respondents report their lives as very satisfying or satisfying [39]. In spite of such ceiling effects, consistent findings across countries lend support to the validity of the measure. Another potential source of bias is the instability of responses due to the mood of the interviewee. As pointed out earlier, this is less of a problem with life satisfaction questions [39] than in other happiness measures. An additional limitation may lie in the exclusion of proxy interviews. Given that cognitive impairment is a major reason for proxy interviews, and that answering life satisfaction questions requires awareness of the interviewee to their current situation, it seems desirable to exclude these individuals with heavy cognitive impairment from our analysis. Finally, concerns could be raised about the subjective health variables. It could be thought these are actually measuring the same latent variable (i.e. overall feeling of health). However, the comparisons between subjective health dimensions are all statistically significant (results available upon request), suggesting that they measure different latent aspects of health. Hence our subjective health measures seem related, but different constructs. Finally. CRELES does not include the times of diagnosis. Therefore, it is not possible to assess whether individuals adapt to particular health conditions over time.

Another concern is the absence of important variables from our analysis. Personality states and traits are particularly important. According to Cheng and Furnham[127], the big five personality factors (openness, contentiousness, extraversion, agreeableness, neuroticism) have consistent relations to happiness. Indeed, neuroticism has been found to be an important happiness correlate in Latin America [44]. However neuroticism, extraversion and openness tend to be stable across time [127-129]. First-differencing and LDVM should have partially washed away the effect of these factors.

We could have claimed causal relationships by removing the endogeneity in our models [111]. The two most important sources of endogeneity are reverse causality (i.e. happiness affecting health), and omitted-variable bias. The FD, LDVM and change-scores analyses reduce the potential of reverse causality. However, there is no certainty that these techniques can control for all relevant omitted variables. Hence, this study cannot claim causality.

This is one of the first analyses that use a nationally representative survey to assess life satisfaction in the elderly. The few studies previously published on elderly happiness in middle-income countries relied often on small-scale surveys [51,52,130] or used only one wave of data [45,131]. Using two waves of data of a nationally representative survey enabled adjusting for the effect of some unobservable factors; the sampling frame of our data allows us to extrapolate the findings to the national population. The quality of the CRELES data set is a further strength. First, the interviewers experienced only 4% of refusals. Second, while the database has some limited missing data; the vast majority of variables have less than 5% missing observations.

Amongst existing health interventions, depression and pain management may be viable instruments to improve the happiness of people over age 60. However, further research is needed to determine which health interventions improve happiness. Subsequent studies could examine which mental conditions are related to happiness. Future research could also examine if appropriate treatment of these conditions leads to happiness. Third, future studies could evaluate the impact on happiness of specific programs for people over age 60. Fourth, the role of adaptation to health conditions needs more thorough research.

Conclusion

Subjective health seems more successful than objective health in improving happiness (measured as life satisfaction). Subjective health in turns depends on a set of determinants, some of which (i.e. socio-economic status or expectations) are not within the direct reach of health care. However, health services can treat effectively some subjective health components, such as pain and depression. Through such interventions, health policy may be more likely to affect happiness.

Tables

Table 1: CRELES Sample, Waves 1 (2004-2006) and 2 (2006-2008)

	Round 1 (n=2,011)	Round 2 (n=1,684)	Diff	p-value
Very satisfied (%)	74.9%	78.2%	3.3%	*0.041*
Age (years)	69.2	70.4	1.1	
Sex (% female)	51.8%	52.5%	0.8%	*0.934*
Education (% no education/ primary school)	76.4%	N/A	N/A	
Civil status (% married or cohabiting)	63.4%	62.1%	-1.3%	*0.472*
Religious frequency (% go to church >=1 time/weekly)	53.3%	55.4%	2.1%	*0.271*
Intense contact with children (% sees children >=1 time/weekly)	79.1%	80.4%	1.4%	*0.284*
Exercises (% exercises >=3 times/week)	34.8%	33.4%	-1.4%	*0.449*
Income (thousands colones, nominal)	132.0	173.1	41.1	*0.000*
Poverty (% below poverty line)	13.8%	12.6%	-1.3%	*0.294*
Assets (score)	13.8	14.3	0.5	*0.112*
Locus of control (score)	3.1	N/A	N/A	*N/A*
Nr. of people in household	3.4	3.2	-0.2	*0.010*
Urban residence (% urban)	63.6%	62.9%	-0.7%	*0.694*
Independent variables				
Subjective health				
Perceived health				
Self-reported health (% with >= Good self-perceived health)	55.2%	60.5%	5.3%	*0.005*
Health compared to peers (% who feel have better health than others)	79.6%	84.2%	4.6%	*0.002*
Mental health				
Depression (% depressed)	6.0%	5.6%	-0.4%	*0.675*
Pain				
Pain in stomach (% with pain)	24.0%	N/A	N/A	*N/A*
Pain lower limbs (% with pain)	50.6%	N/A	N/A	*N/A*
Objective health				
Disability (% with >=+5 ADL/IADL)	10.2%	10.6%	0.4%	*0.715*
Chronics (% with >=3 chronic conditions)	28.2%	29.1%	0.9%	*0.586*

Allostatic index (% with low AL)	35.3%	47.3%	12.0%	*0.000*
Metabolic syndrome (% with metabolic syndrome, IDF criteria)	47.6%	56.9%	9.2%	*0.000*

** Analyses take probability sampling weights into account*
** Education and locus of control available only in wave 1*

Table 2: Relationship between Life Satisfaction and Independent Variables, Rounds 1 and 2

	Round 1 (n=2,011)			Round 2 (n=1,684)		
	Very satisfied	**< Very satisfied**	**p-value**	**Very satisfied**	**< Very satisfied**	**p-value**
Subjective health						
Perceived health						
Self-reported health (% with >= Good self-perceived health)	63.6%	30.2%	*0.000*	68.5%	31.7%	*0.000*
Health compared to peers (% who feel have better health than peers)	84.3%	65.3%	*0.000*	87.6%	71.6%	*0.000*
Mental health						
Depression (% depressed)	0.9%	21.1%	*0.000*	1.1%	21.7%	*0.000*
Pain						
Pain in stomach (% with pain)	19.2%	36.7%	*0.000*	N/A	N/A	*N/A*
Pain lower limbs (% with pain)	45.1%	65.3%	*0.000*	N/A	N/A	*N/A*
Objective health						
Disability (% with >=+5 ADL/IADL)	8.1%	16.6%	*0.000*	8.3%	19.0%	*0.000*
Chronics (% with >=3 chronic conditions)	25.0%	37.7%	*0.000*	26.9%	37.2%	*0.001*
Allostatic index (% with low AL)	35.5%	34.6%	*0.180*	49.3%	40.5%	*0.015*
Metabolic syndrome (% with metabolic syndrome, IDF criteria)	46.6%	50.6%	*0.750*	53.6%	68.7%	*0.000*

** Analyses take probability sampling weights into account*

Table 3. Main analysis: First-Differences Model for Life Satisfaction (predicted probabilities)

	Predicted Probabilities					
	Subjective Health				Objective Health	
	Self-reported health	**Std. Error**	**Depression**	**Std. Error**	**Disability**	**Std. Error**
Persistent dissatisfaction	-4.3%	*0.005****	11.6%	*0.012****	1.3%	*0.007**
Satisfaction decrease	-2.4%	*0.003****	6.4%	*0.009****	0.7%	*0.004**
Satisfaction increase	-2.0%	*0.003****	5.6%	*0.008****	0.6%	*0.003**
Persistent satisfaction	8.7%	*0.087****	-23.6%	*0.024****	-2.6%	*0.014**

** Outcome is modeled as follows: persistent dissatisfaction (< very satisfied at both time points); satisfaction decrease (very satisfied - < very satisfied); satisfaction increase (<very satisfied - very satisfied); persistent satisfaction (very satisfied at both timepoints).*
** Control covariates include subjective health (self-rated health, depression), objective health (disability, chronic diseases, allostatic index, metabolic syndrome), socio-economic measures (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), and cognitive impairment*
** Results reported only for statistically significant variables*
** Results express predicted probabilities for 1-unit change in independent variables*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 4b. Main Analysis: First-Differences Model for Life Satisfaction (standardized coefficients)

	Standardized	Std. Error
Subjective health		
Self-reported health	0.296	*0.033****
Depression	-0.362	*0.085****
Objective health		
Disability (+5 ADL/IADL)	-0.048	*0.048****

** Control covariates are same as in table 3 (first-differences model covariates)*
** Results reported only for statistically significant variables*
** Results express standard deviation in Y per standard deviation change in X*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 5. Sensitivity Analysis - Lagged Dependent Variable Model (LDVM) for Life Satisfaction

	Standardized	Marginal	Std. Error
Subjective health			
Self-reported health	0.240	16.9%	*0.026****
Depression	-0.242	-50.9%	*0.081****
Objective health			
Disability (+5 ADL/IADL)	-0.048	-5.2%	*0.038*
Disability ADL	-0.081	-5.5%	*0.025***
Disability IADL	-0.049	-4.2%	*0.031*
Chronics (three or more)	0.012	0.9%	*0.025*
Allostatic index	0.012	-1.2%	*0.023*
Metabolic syndrome	-0.076	-5.0%	*0.024***

** Control covariates include subjective health measures (self-rated health, depression), objective health (disability, chronic diseases, allostatic index, metabolic syndrome), socio-economic measures (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), cognitive impairment, and baseline life satisfaction.*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 6. Sensitivity Analysis – Pooled Cross-Sectional Model for Life Satisfaction

	Standardized	Marginal	Std. Error
Subjective health			
Self-reported health	0.255	0.181	*0.016****
Depression	-0.323	-0.596	*0.036****
Objective health			
Disability (+5 ADL/IADL)	-0.060	-0.066	*0.025***
Chronics (>=3)	-0.027	-0.020	*0.018*
Allostatic index	0.001	-0.001	*0.017*
Metabolic syndrome	-0.016	-0.011	*0.017*

** Control covariates include subjective health measures (self-rated health, depression), objective health (disability, chronic diseases, allostatic index, metabolic syndrome), socio-economic (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), and cognitive impairment*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 7a. Sensitivity Analysis – Change-scores Model for Life Satisfaction (Predicted Probabilities)

ΔLife Satisfaction	Predicted Probabilities									
	Subjective Health								Objective Health	
	Self-reported health	Std. Error	Depression	Std. Error	Pain stomach	Std. Error	Pain lower limbs	Std. Error	Disability	Std. Error
Persistent dissatisfaction	-7.1%	*0.017****	20.8%	*0.024****	3.3%	*0.013***	2.5%	*0.012***	3.9%	*0.018***
Satisfaction decrease	-4.1%	*0.015****	2.5%	*0.024*	1.9%	*0.008***	1.4%	*0.007***	2.3%	*0.011***
Satisfaction increase	-6.9%	*0.015****	24.3%	*0.058****	0.9%	*0.004***	0.7%	*0.003***	1.1%	*0.005***
Persistent satisfaction	18.1%	*0.022****	-47.5%	*0.062****	-6.1%	*0.024***	-4.6%	*0.022***	-7.3%	*0.034***

** Control covariates include subjective health measures (self-rated health, depression, pain stomach, pain lower limbs), objective health (disability, chronic diseases, allostatic index, metabolic syndrome), socio-economic (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), and cognitive impairment*

** Results express predicted probabilities for 1-unit change in independent variables*

** Only statistically significant subjective and objective health measures are reported*

** Robust standard errors*

** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 8b. Sensitivity Analysis – Change-Scores Model for Life Satisfaction (Standardized Coefficients)

	Standardized	Std. Error
Subjective health		
Self-reported health	0.229	*0.076****
Depression	-0.241	*0.14****
Pain stomach	-0.073	*0.080***
Pain lower limbs	-0.068	*0.072***
Objective health		
Disability (+5 ADL/IADL)	-0.061	*0.114***

** Same control variables as in table 6a*
** Only statistically significant subjective and objective health measures are reported*
** Results express standard deviation change in Y per standard deviation change in X*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 9. First-Differences Model for Life Satisfaction by Disability Subtype

	Disability Sub-Types (ADL and IADL)			
	Predicted Probabilities			
	ADL	**Std. Error**	**IADL**	**Std. Error**
Persistent dissatisfaction	0.010	*0.005**	0.005	*0.006*
Satisfaction decrease	0.006	*0.003**	0.003	*0.003*
Satisfaction increase	0.005	*0.003**	0.002	*0.003*
Persistent satisfaction	-0.021	*0.011**	-0.011	*0.012*

** Outcome is modeled as follows: persistent dissatisfaction (< very satisfied at both time points); satisfaction decrease (very satisfied - < very satisfied); satisfaction increase (<very satisfied - very satisfied); persistent satisfaction (very satisfied at both timepoints).*
** Models use first-differences covariates: subjective health (self-rated health, depression), objective health (disability, chronic diseases, allostatic index, metabolic syndrome), socio-economic (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), and cognitive impairment*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 10. Lagged Dependent Variable Model for Life Satisfaction by Disability Subtypes

	Standardized	**Marginal**	**Std. Error**
Subjective health			
Self-reported health	0.240	0.169	*0.026****
Disability sub-types			
Disability ADL	-0.081	-0.055	*0.025***
Disability IADL	-0.049	-0.042	*0.031*

** Models use LDVM covariates: subjective health (self-rated health, depression), objective health (disability type, chronic diseases, allostatic index, metabolic syndrome), socio-economic (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), cognitive impairment, and baseline life satisfaction.*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 11. Mediation Tests Self-reported Health (LDVM Models)

	Total effect	Std. Error
Mediation between depression and life satisfaction		
Total effect	-0.421	*.032****
Direct effect of depression on life satisfaction	-0.363	*.035****
Effect mediated by self-reported health	-0.058	*.0134****
% of total effect mediated by self-reported health	13.3%	

	Total effect	Std. Error
Mediation between poverty and life satisfaction (n=278)		
Total effect	-0.046	*0.042*
Direct effect of poverty on life satisfaction	-0.045	*0.041*
Effect mediated by self-reported health	-0.001	*0.009*
% of total effect mediated by self-reported health	-0.1%	

** Covariates are: subjective health (self-rated health, depression), objective health (disability type, chronic diseases, allostatic index, metabolic syndrome), socio-economic (age, sex, education, civil status, poverty status, household assets, number of household members), lifestyle (attending church, exercising, children), cognitive impairment, and baseline life satisfaction*
** Model regresses happiness at wave 2 on wave2 covariates and baseline life satisfaction*
** Standard errors are robust*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Appendix: additional estimations

Perceived health measure: self-assessment of own health status versus peers

Table A1: Fist differences model (predicted probabilities)

	Predicted Probabilities					
	Subjective Health				Objective Health	
	Health compared to peers	**Std. Error**	**Depression**	**Std. Error**	**Disability**	**Std. Error**
Persistent dissatisfaction	-3.3%	*0.007****	13.2%	*0.013****	NS	NS
Satisfaction decrease	-1.9%	*0.004****	7.6%	*0.011****	NS	NS
Satisfaction increase	-1.8%	*0.004****	7.0%	*0.010****	NS	NS
Persistent satisfaction	7.0%	*0.014****	-27.8%	*0.027****	NS	NS

** Same covariates as for table 3*
** Outcome is modeled as follows: persistent dissatisfaction (< very satisfied at both time points); satisfaction decrease (very satisfied – < very satisfied); satisfaction increase (<very satisfied – very satisfied); persistent satisfaction (very satisfied at both timepoints).*
** Results reported only for statistically significant variables*
** Results express predicted probabilities for 1-unit change in independent variables*
** Robust standard errors*
** NS denotes lack of statistical significance*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table A2: First Differences Model (standardized coefficients)

	Standardized	Std. Error
Subjective health		
Health compared to peers	0.166	*0.047****
Depression	-0.412	*0.093****
Objective health		
Disability (+5 ADL/IADL)	N/S	*N/S*

** Same control covariates as in table 3 (first-differences model covariates)*
** Results reported only for statistically significant variables*
** Results express standard deviation in Y per standard deviation change in X*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table A3: Lagged Dependent Variable Model

	Standardized	Marginal	Std. Error
Subjective health			
Health compared to peers	0.087	8.8%	*0.036***
Depression	-0.270	-55.5%	*0.074****
Objective health			
Disability (+5 ADL/IADL)	-0.050	-5.5%	*0.040*
Disability ADL	-0.119	-7.9%	*0.024***
Disability IADL	-0.063	-5.5%	*0.033**
Chronics (three or more)	-0.008	-0.5%	*0.026*
Allostatic index	0.010	0.7%	*0.024*
Metabolic syndrome	-0.048	-3.2%	*0.025*

** Outcome: probability of feeling very satisfied*
** Control covariates are same as in table 4.*
** Robust standard errors*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

3. THE RELATIONSHIP BETWEEN EARLY CHILDHOOD HEALTH AND HAPPINESS IN PEOPLE OVER AGE 60

Abstract

The determinants of happiness in senior populations are fairly well understood. However, the link between childhood health status and senior happiness has not yet been determined. We investigate whether childhood health is associated with happiness (measured as life satisfaction) at later age, and the potential pathways for such association. Our study uses a two-wave, nationally representative (n=2,827) survey of the elderly in Costa Rica (CRELES). Cross-sectional and lagged dependent variable models support a statistically significant and meaningful association between poor childhood health and elderly life satisfaction (β=-0.071; p=0.06). Mediation tests suggest that adult self-reported health may be the main mediator in this association. Child health may be positively associated with life satisfaction throughout the life course. Further research can elucidate the pathways for such a connection.

Introduction

Happiness is an ultimate human pursuit and a majority of people share happiness as a life goal[132]. In an international survey of 7000 respondents in 42 countries, 69% of rated happiness as their top priority[60]. Mounting evidence also suggests that happiness may impact health. For example, research has shown that optimistic people tend to survive longer, due to improved self-efficacy and possibly physiological effects[14]. Indeed, happiness is a necessary component of health according to its most common definition, which emphasizes "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity”[35].

Happiness research is new and rapidly evolving. A majority of the literature defines happiness as life satisfaction, that is “[the factors that] lead people to evaluate their life in positive terms” [85]. The determinants of what makes an adult happy have been well studied over the past 15 years, and several factors emerge as primary determinants. Genetics explains a large component[37] of the variation in happiness across adults. Personality traits such as neuroticism[133,134], extraversion[133,134], and locus of control[75, 76] are consistently associated with happiness. While genetics and personality traits seem to remain with the person for life, other predictors such as health, income, civil status, employment, and social ties can change over the lifetime[38-40]. Additional dimensions correlated with adult happiness include age[38,62] religious beliefs[135] and environmental factors, such as safety, urbanization[38] and the political system[69]. Among these time-variant factors, health status stands out [41,42,43,44] as particularly important among the elderly[45].

Less research has investigated the links between childhood health status and adult happiness. Related fields underline the importance of early childhood: strong links have been found between early life conditions (i.e. childhood health[56,136, 137, 138,136,139]; socioeconomic status [57,140,56,141]) and adult health. We investigate if childhood health could be related to happiness in late life. So far, this question was addressed by only one study[58], without significant results. However, the previous effort relied exclusively on objective childhood health measures (disability and chronic conditions), and did not explore causation channels.

This study examines happiness in the senior population of a middle income country- Costa Rica. It takes advantage of a data set that measures subjective and objective happiness in seniors, and contains information on their subjective childhood health status. The dataset, patterned after the Health and Retirement Survey of the US, contains a nationally representative sample of seniors.

This effort examines two interconnected null hypotheses. The first is that childhood health is not associated with life satisfaction (i.e. happiness) at a later age, adjusting for relevant confounders. The second postulates that adult health does not mediate the previous relationship, adjusting for the same confounders. The policy implication is that attention given to child health can impact happiness throughout the life course. Answering these questions is particularly important to Latin America and the Caribbean, a region projected to age rapidly[59]; however, it has implications for the US and other high income countries as well.

Methods

Database

This study uses the "Costa Rican Study on Longevity and Healthy Aging" (CRELES)[81]. CRELES is a nationally representative sample of residents of Costa Rica aged 60 years and older, which covers data on self-reported physical health, psychological health, living conditions, health behavior, health care utilization, social support, and socioeconomic status. This survey conducted two rounds of interviews on the same individuals (n=2,827 and n=2,364 respectively), separated by two and a half years. Three features of the CRELES data are relevant for analytical purposes. First, the difference in sample sizes between both waves indicates that the attrition rate was 16% (n=463). A majority (10%) of non-respondents were deceased by the second interview, and the remainder (6%) could not be located, mainly due to incorrect address[82]. Bivariate analyses revealed that censored and non-censored individuals were not statistically different in terms of life satisfaction, socio-demographic, and health. Hence, attrition-generated bias should be modest. Second, nonresponses at baseline were low: only 4% of interview candidates refused to be interviewed at the first wave [81]. Finally, CRELES oversampled the oldest-old (+95 years old). The survey incorporates a set of weights to correct for oversampling and non-response bias[83].

At CRELES, 703 interviews in the first and 676 in the second wave used a proxy respondent [73], due to the senior's inability to communicate and/or cognitive impairment as measured by the Mini-Mental Measurement Scale (MMSE). These interviews do not include any self-reported variables, including life satisfaction. Proxy interviewees were therefore excluded, which results in an analytical sample of n=2,111 (wave1) and n= 1,684 (wave 2). As mentioned in the previous chapter, the exclusion of these individuals is desirable for analytic purposes.

Conceptual framework

Our model[39] regresses senior life satisfaction on childhood health, other childhood conditions (childhood poverty, single-parent home), and a set of control covariates selected from the happiness literature[38,39]. Control covariates include socio-economic variables (age, gender, education, and others); social factors (seeing adult children, frequency of attendance to religious services, etc.); health status (subjective and objective); and personality characteristics (locus of control). The following sections explain these in more detail.

Variables

This study uses life satisfaction as the dependent variable. There are many definitions of happiness, but life satisfaction is considered more stable over time than competing definitions such as positive attitude, affect, and morale[39]. The dependent variable was extracted from the most commonly used happiness question[89,12]: "in general, how do you feel about your life?: very satisfied, somewhat satisfied, somewhat unsatisfied, very unsatisfied". The same question is asked in surveys in the US[36,142], Germany[67], and Europe-wide[88] surveys. As shown in table 1, three-fourths of respondents in both waves reported feeling very satisfied. Under such conditions, power calculations determined that variability in some categories of the outcome would not be sufficient for statistical inference. Therefore, the dependent variable was dichotomized into being very satisfied or less than very satisfied.

All childhood questions in CRELES refer to the first 15 years of life (i.e. childhood health or early life). The main independent variable (i.e. poor childhood health) is self-reported from the following question: How was your health for the majority of your childhood and adolescence? - Excellent, Very Good, Good, Poor. Given it is self-reported, the main outcome is a subjective measure of health. Secondary measures of childhood health concern specific conditions: tuberculosis, rheumatic fever, poliomyelitis, malaria and asthma/chronic bronchitis. These, in contrast to the primary outcome, are objective health measures.

The model controls also for childhood adversity (poverty and father absenteeism). Overall childhood poverty is determined by an affirmative response to having suffered from economic hardship during early life, which resulted in regular inadequacy of eating, dressing or medical care (Yes/No). Finer questions on poverty include not having worn shoes, having lived in a home without a bathroom or latrine, and having lived in a home without electricity during early life. The one-parent household variable records whether the biological father was absent from the childhood home.

CRELES includes assessments of subjective and objective health. Subjective health measures concern typically non-physical health states, such as perceived health and depression[143]. At CRELES, four questions ask respondents to rate their own health. Two were asked in both waves: self-reported health (=1 if the interviewee reported feeling in good, very good or excellent health) and feeling in better health than others (=1 if the respondent feels in better health than others their age). The other two are only available in the first wave. These are health status selected from a card (=1 if self-rated their health >=5 from a card, where 1=poor health and 7=excellent health) and longevity (=1 if expects to live >= 10 more years; question asked to respondents up to 85 years old). All these variables were also dichotomized along the good vs. fair/poor axis. Individuals are classified as depressed if they suffer 10 or more depression symptoms on the Yesavage Geriatric Depression Scale (GDS-15) [91].

Objective health measures are disability, chronic diseases, allostatic index, and metabolic syndrome. These are dichotomized from index variables, which were developed from measurements taken during the interviews. Disability expresses inability to perform at least 5 out of 14 Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL)[92]. The metabolic syndrome variable follows the criteria of the International Diabetes Federation[94]. The chronic disease variable denotes those with 3 or more chronic diseases as measured by a count of self-reported diagnosis of cancer, heart attack, heart disease, stroke, hypertension, high blood lipids, diabetes, lung disease, arthritis, and osteoporosis. Allostatic load represents the ‘wear and tear’ of the body and is a sum of seven biomarkers. Following convention, a point is awarded for each dimension if the concerned individual is in the top 25 percentile of his or her distribution [96].

Additional covariates are age (5-year age groups); gender; civil status (=1 if married or cohabiting), educational level (4 categories: no education, primary, secondary, tertiary); number of household members (continuous); urban residence (dichotomous); health region (7 health regions); living under the poverty line; housing assets; and locus of control. Locus of control sums 7 items reflecting whether respondents feel in control of their lives [98]. Individuals were considered poor if their personal income was below the national poverty line. Personal income had already been adjusted to the number of people living in the home and whether the respondent lived in a home that they own. Finally, the housing assets index is a weighted average (range 0–32) of housing quality (quality and materials of floor, ceiling) and home assets (number of telephones, of computers, and others) [102].

Imputation of missing data

At CRELES, missing data is not a significant problem: only 5% of observations were missing two or more happiness predictors. Variables with missing data were imputed. Three covariates (depression scale, locus of control, and allostatic index) were missing in >5% of observations. Imputation proceeded through two different methods. First, single imputation proceeded for index variables (i.e. variables that are the sum of subvariables). These are depression, disability, locus of control and assets. Single imputation was appropriate as information was available on the majority of subvariables in each index (only <5% of observations were missing >20% of items in the index). The rest of variables (chronic diseases, allostatic index, frequency of seeing children, metabolic syndrome) were imputed through multiple imputation chained equations (MICE)[104]. MICE predicts missing values based on other variables with non-missing values; it imputes first the variable with the least number of missing observations; then the second, and so on. All variables are cycled through in such manner, and the whole process is repeated (usually 5-10 times) until a dataset with converged estimates has been reached. This process produces a single imputed dataset with no missing data[105].

Statistical analysis

Exploratory data analyses explored first the late life differences between those who had suffered poor versus good childhood health. Then, bivariate analyses explored the relationships between childhood adversity and senior life satisfaction. Finally, the multivariate relationships between childhood health and senior life satisfaction were explored.

The main analysis of this study consists in a lagged dependent variable model (LDVM). LDVM models were used for estimating the associations with life satisfaction of both the primary (overall childhood health) and secondary independent variables (specific childhood conditions). LDVM can partially remove endogeneity. Endogeneity, which results mainly from mutual causation and omitted variable bias[106], can pose a serious threat to statistical inference[111]. By including lagged life satisfaction, LDVM models adjust for mutual causation[110], and partially for unobservable historical confounders[106], which include personality characteristics [144], and other life circumstances. The estimates of LDVM models are reported as marginal effects. These measure the probability change in the outcome associated with a one unit increase in the independent variable.

LDVM regresses second round life satisfaction on childhood health, round two predictors and baseline life satisfaction.

$$Probit(Life\ Satisfaction_2)= \beta_0+ \beta_1(Childhood\ health)+ \beta_2(Childhood\ poverty)+ \beta_3(Single\text{-}parent\ childhood\ home)+\beta_4(Age_2)+ \beta_5(Sex_2)+ \beta_6(Civil\ Status_2)+ \beta_7(Household\ residents_2) +\beta_8(Area\ of\ residence_2) + \beta_9(Attends\ Church_2)+ \beta_{10}(Sees\ Children_2)+ + \beta_{11}(Poor_2)+ \beta_{12}(Assets_2)+ \beta_{13}(Life\ Satisfaction_1)+\varepsilon$$

Sensitivity analyses are pooled cross-sectional and change-scores. Both focus on the primary independent variable only (overall childhood health) and use the same controls as the main analyses. Pooling both rounds of data increases sample size, and hence estimate precision, versus running separate cross-sections. The pooled cross-sectional model controls for temporal effects through a year dummy variable. However, as any cross-sectional model, it is highly vulnerable to omitted variable bias [111].

Change-scores models are an alternative to LDVM; however the econometric literature considers them to be more prone to estimation problems [110]. Our change-scores model regresses the changes in life satisfaction on baseline predictors (ΔY on X_1). Following the approach from Roberto and Gaskin [107], changes in happiness adopt the following form: 0 for persistent dissatisfaction (less than very satisfied at both time points), 1 if satisfaction decreased between both waves, 2 if satisfaction increased between both waves, and 3 for persistent satisfaction (i.e. individuals very satisfied at both waves). Hence, these are ordered from the least to the most desirable outcome. Given that the dependent variable is categorical with four levels, this model uses an ordered probit specification, and expresses the results as predicted probabilities. These provide the probability change in a given level of the outcome, resulting from a one unit-increase in the independent variable. Hence, each level of the outcome will have its own predicted probability.

Interactions: does the association between childhood health status and senior life satisfaction vary across childhood adversities?

Interactions explore whether the relationship between childhood health and elderly life satisfaction varies across other forms of childhood adversity. These interactions use wave 1 data, and adopt the following form:

$$Probit(Life\ Satisfaction_1) = \beta_0 + \beta_1\ (Childhood\ health*Childhood\ condition) + \beta_2\ (Childhood\ health) + \beta_3 (Childhood\ condition) + \beta_4 (Childhood\ poverty) + \beta_5 (Single\text{-}parent\ home) + \beta_6 (Age_1) + \beta_7\ (Sex) + \beta_8 (Civil\ status_1) + \beta_9\ (Household\ residents_1) + \beta_{10}\ (Area\ of\ residence_1) + \beta_{11} (Attends\ Church_1) + \beta_{12} (Sees\ Children_1) + \beta_{13} (Poor_1) + \beta_{14} (Assets_1) + \varepsilon$$

The general interaction is expressed by (Childhood health*Childhood condition). The first term of the interaction is always childhood health, which was interacted with 5 different childhood conditions: overall poverty; having lived in a home with no bathroom as a child; having lived in a home with no electricity; have not worn shoes regularly; and one-parent household.

Mediator variables: potential vehicles for the relationship between childhood health status and senior happiness

Assuming there is an association between childhood health and adult life satisfaction (hypothesis 1), the next question is the nexus for such relationship (hypothesis 2). The most likely links are health (subjective and objective) and personality variables. In statistics, mediators stand in the causal chain and affect the direction and magnitude of relationships between the independent and dependent variables[113]. The availability of both rounds of data for some variables (self-reported heath; health compared to peers; functional disability; chronic diseases and depression) permitted testing their mediation through LDVM models. Other variables (health status as chosen from a card; expected longevity; pain and locus of control) were available only in the first survey round; their mediation was therefore tested through cross-sectional models. As discussed earlier, cross-sectional models are less methodologically robust, given they are more vulnerable to endogeneity (i.e. reverse causality and omitted variable bias) than LDVM.

In all specifications, mediation was tested using a specialized command (binary_mediation) in Stata 11[112]. Mediation tests used a full happiness model:

$$Probit(Life\ Satisfaction_2)=\beta_0+ \beta_1\ (Mediator_2)+ \beta_2\ (Childhood\ health)+ \beta_3(Childhood\ poverty)+ \beta_4(Single\text{-}parent\ childhood\ home)+ \beta_5\ (Age_2)+ \beta_6\ (Sex)+ \beta_7\ (Civil\ status_2)+ \beta_8\ (Nr.\ Household\ residents_2)\ +\beta_9\ (Area\ of\ residence_2)\ + \beta_{10}(Attends\ Church_2)+ \beta_{11}(Sees\ Children_2)+ \beta_{12}(Poor_2)+ \beta_{13}(Assets_2)+ \beta_{14}(Health\ covariates_2)+ \beta_{15}(Personality\ covariates_2)+ \beta_{16}(Life\ Satisfaction\ _1)+\ \varepsilon$$

A note of caution is warranted on the mediation test: standard mediation tests, such as the one used in this essay, assume that the mediator (i.e. perceived health) is randomly assigned[113], and hence unaffected by unobservable confounders. However, health and personality variables are likely to be influenced by dimensions not captured by our model, which would violate such assumption. Therefore, the results from the mediation tests must be interpreted with caution.

Falsification variables

The same models were used to test falsification variables. These are important happiness predictors that should not mediate between childhood health and senior life satisfaction. Falsification dimensions are attending religious services, seeing children frequently, urban residence, and poverty. Insignificant mediation tests would suggest that childhood health is not associated with life satisfaction through these dimensions.

Additional measures were used to enhance statistical estimation. These include adjusting for heteroskedasticity through robust standard errors, and assessment of collinearity through variance inflation factors.

Results

The multivariate analyses support that childhood health status is significantly associated with life satisfaction in people over 60. This finding is also buttressed by the bivariate relationships in our data, according to which a) those who suffered poor childhood health are worse off across multiple dimensions at late age, and b) poor childhood health and adversity are associated with less life satisfaction after age 60.

Table 1 compares the characteristics of those who experienced good and poor childhood health. After age 60, those in poor childhood health (all differences $p<0.05$) are less satisfied with their lives, younger, less educated, reside in the country side, have lower personal incomes and less housing assets. They experienced also more childhood adversity: disease, poverty, and family dysfunction (single-parent homes where the mother had no formal education). Hence, poor childhood health is related with social disadvantage, in both early and late life.

Table 2 shows that poor childhood health and adversity are associated with more unhappiness after age 60. In the less satisfied group, 12.8% had suffered ill overall health as children, almost double as much as in the more satisfied group (6.9%; $p<0.05$). Less satisfied individuals were twice as likely to have suffered certain childhood diseases (malaria and asthma ($p<0.05$)). Finally, less satisfied seniors experienced also more childhood poverty in three different measures ($p<0.05$), and higher (non-significant) probabilities of having been raised in a single-parent household.

Table 3 displays the bivariate relationship between childhood health status and mediator variables. As discussed in the methods section, potential mediators are current subjective and objective health. Significant relationships are found only for subjective health mediators. Within these, childhood health status is most strongly associated with adult self-reported health (14% group difference; $p=0.003$). Poor childhood health is also associated with a considerable higher probability of clinical depression (8.7%; $p<0.000$). In contrast, childhood health status did not exhibit any significant relationships with the objective health mediators. These results suggest strongly that childhood health is related to how people *feel* about their late life health, rather than how healthy they really *are*. In consequence, subjective health—particularly self-reported health—may be the main link between childhood health and life satisfaction after age 60.

Table 4 shows the multivariate regressions of senior life satisfaction on childhood health. In the main analysis (LDVM model), poor childhood health (primary outcome) is associated with a 7.1% lower probability of being very satisfied ($p<0.06$) at wave 2. In specific conditions, only childhood asthma bears a relationship with senior life satisfaction (-6.8%; $p<0.08$). This suggests that overall childhood health is a more reliable predictor of senior happiness than objective (disease-specific) measures.

Both sensitivity analyses (pooled cross-sectional and change-scores; also in table 4) confirm that overall childhood health has a negative significant association with senior happiness. The consistency of these estimates rejects hypothesis 1, and therefore supports a relationship between childhood health and elderly happiness. Further, table 5 shows that none of the interactions between childhood health and other childhood adversity dimensions (not having worn shoes, and so on) were statistically significant. Hence, the data support an association between childhood health and senior happiness, but this relationship is independent of other childhood adversities.

The mediation tests in Table 6 lead us to reject hypothesis 2, as they find that self-reported health and locus of control mediate (respectively) 31.4% ($p<0.1$) and 18.3% ($p<0.05$) of the relationship between childhood health and life satisfaction. Adult-self reported health is positively correlated with both childhood health and adult life satisfaction. Hence, lower childhood health is associated with lower adult health, which translates into lower life satisfaction. This mediation was tested through an LDVM model. However, locus of control is negatively related to childhood health (while still being positively related to life satisfaction). The negative association between childhood health and locus of control accounts runs contrary to theory: better childhood health is associated with lower perception of control, and hence less life satisfaction. However, this test used cross-sectional data only (as locus of control was collected only in the first wave); omitted-variable bias may account for the unexpected direction in this mediation. Taken together, these facts support that adult self-perceived health is the main mediator of our relationship of interest.

Discussion

This study finds that childhood health is associated with senior life satisfaction (i.e happiness), and that adult self-reported health may be the main mediator for such association.

All models support statistically significant and important associations between childhood health and senior life satisfaction. In the LDVM model, childhood health (β=-0.07; p=0.06) is a strong a predictor as weekly church attendance (β=-0.07; p=0.001) - a factor long associated with happiness in older adults. Childhood health is more important that other conventional predictors, such as civil status (β =0.04; p=0.07) and adult poverty (β=-0.04; p=0.21)). In addition, the self-perceived measure of overall childhood health is a stronger predictor of senior life satisfaction than the disease specific measures. Finally, mediation analyses support that adult self-perceived health mediates 31.4% of the positive association between childhood health and senior life satisfaction.

The connection between childhood health and adult life satisfaction could take place through pathway, latency and cumulative effects. All these have been identified for the relationships between child and adult health[145,138]. Pathway effects "occur when early-life circumstances influence life trajectories, creating indirect effects on future health"[145]. For example, childhood malnutrition could create learning difficulties; the resulting lower educational levels could affect health and hence life satisfaction. Latency effects consist of the "independent effects [on future health] of early-life exposures". In such case, poor childhood health would create a persistent, pessimistic outlook early in life that persist throughout the life course. Finally, cumulative effects occur when health in late life is affected by repeated stressors over the life course" [145]. In such case, the same factors (i.e. poverty, living in a marginalized area, etc.) that would affect our mediators also life satisfaction throughout the life course.

The current study has some limitations. First, some questions in the CRELES questionnaires may be prone to recall bias. For example, the childhood health question inquired about the respondent's health during "the majority of childhood and adolescence." Memory gives leeway to recall biases[146]; assessing the "majority of childhood and adolescence" presents cognitive challenges. While recognizing these limitations, it must be pointed out that other major surveys are similarly worded[147,148]. Second, there is no guarantee that our analysis was able to control for all relevant confounders, particularly time-invariant confounders (i.e. genetics), which are crucial in happiness research[80]. Another limitation lies in availability of only one wave of data for some mediators (good perceived health; longevity; locus of control). These could be only assessed through cross-sectional models; the resulting omitted-variable bias may have biased our findings. Finally, the absence of life course data makes it impossible to assess association pathways across life stages.

These limitations do not mitigate the study's strengths. First, previous literature on elderly happiness relies usually on smaller surveys [130,51,52,149]. Hence, our study overcomes previous concerns of low statistical power and the impossibility to generalize to the national scale. Second, the few studies that do use large datasets[45,131] rely primarily on cross-sectional data. By using two waves of data, we adjusted for unobservable historical confounders through LDVM[106]. Finally, LDVM reduces the possibility of mutual causation[110]. It is possible that unhappy individuals tend to artificially over report adverse conditions, including their childhood health. In such case, mutual causation would occur, i.e. exposure and outcome would cause each other. By including the lagged dependent variable, the possibility of life satisfaction impacting reports of childhood health is adjusted for.

The current study highlights areas for further research. Future studies can explore more extensively the association of other childhood dimensions (i.e. family structure, characteristics and relationship with parents, etc.) with happiness. At best, these studies would use semi-experimental designs (instrumental variable or similar) to assess causality. Life course studies could also elaborate on the mediators suggested in this study, particularly on the role of locus of control. Finally, further research could explore other potential mediators. Such body of research could form the empirical basis to incorporate childhood into existing happiness theory.

Conclusion

Our results suggest that childhood health status is an important predictor of life satisfaction (a measure of happiness) in adults over age 60. Thus, childhood health may be one of the best possible happiness investments[150,14]. There are cost-effective interventions to improve childhood health, including neonatal care packages[151], immunizations, and fortifying food through vitamin A and zinc[152]. Through such interventions, governments may improve both the health and the wellbeing of individuals over the life course. Further light into such relationships may be a critical addition to the field of happiness research.

Table 1: Late Life Differences by Childhood Health Status, Round 1

	Good childhood health (n=1,929)	Poor childhood health (n=176)	Difference	P-value of difference
Very satisfied (%)	76.3%	60.5%	15.8%	*0.000*
Age (years)	69.3	67.9	140.0%	*0.010*
Sex (% female)	51.7%	52.0%	-0.3%	*0.941*
Married/cohabiting (%)	63.5%	62.1%	1.4%	
Years of education	5.7	4.6	1.1	*0.003*
Nr. household members (count)	3.4	3.4	0.0	*0.765*
Urban residence (%)	64.9%	51.5%	13.4%	*0.003*
Average income (thousand colones)	136.1	95.3	40.8	*0.007*
Below poverty line (%)	13.4%	18.6%	-5.2%	*0.096*
Assets (points)	14.0	12.5	147.1%	*0.028*
Locus of control (points)	3.1	3.2	-3.6%	*0.322*
Educational level of the mother (% mothers with no education)	27.4%	42.1%	-14.8%	*0.001*
Childhood variables				
Childhood poverty (%)	55.4%	70.5%	-15.1%	*0.003*
Did not wear shoes regularly as child (%)	59.1%	74.8%	-15.8%	*0.001*
No electricity in childhood home (%)	62.6%	75.6%	-13.1%	*0.008*
No bathroom/latrines in childhood home (%)	23.9%	37.3%	-13.4%	*0.001*
Father absent from childhood home (%)	19.7%	34.2%	-14.5%	*0.000*
Childhood health variables				
Tuberculosis (%)	0.3%	1.3%	-1.0%	*0.110*
Rheumatic fever (%)	1.2%	8.2%	-7.1%	*0.000*
Poliomyelitis (%)	0.5%	0.0%	0.5%	*0.518*
Malaria (%)	9.4%	26.6%	-17.2%	*0.000*
Asthma/chronic bronchitis (%)	9.3%	28.7%	-19.4%	*0.000*

** Analysis takes sampling weights into account*
** Statistical significance tests are two-tailed tests of proportions (categorical variables) and t-tests (continuous variables)*

Table 2: Relationship between Senior Life Satisfaction and Conditions Experienced during Childhood

CRELES, wave 1

	Very satisfied	**< Very satisfied**	**Difference**	**Test of differences**
Poor overall child health	7.2%	14.0%	6.8%	*0.000*
Childhood poverty (%)	54.6%	63.2%	8.5%	*0.004*
Did not wear shoes regularly as child (%)	58.8%	65.4%	6.6%	*0.026*
No electricity in childhood home (%)	61.9%	69.4%	7.5%	*0.012*
No bathroom/latrines in childhood home (%)	23.8%	29.1%	5.3%	*0.033*
Single-parent home (%)	20.1%	23.7%	3.6%	*0.125*
Childhood health variables				
Tuberculosis (%)	0.4%	0.4%	-0.1%	*0.859*
Rheumatic fever (%)	1.5%	2.7%	1.3%	*0.107*
Poliomyelitis (%)	0.5%	0.5%	0.0%	*1.000*
Malaria (%)	9.8%	14.6%	4.8%	*0.006*
Asthma/chronic bronchitis (%)	9.6%	15.2%	5.6%	*0.003*

** Analysis takes sampling weights into account*
** Statistical significance tests are two-tailed tests of proportions (categorical variables) and t-tests (continuous variables)*
** Sample size too small for statistical tests of differences for tuberculosis and poliomyelitis.*

Table 3: Bivariate Relationships of Childhood Health with Mediators and Falsification Variables

CRELES, Round 1

Mediators	**Good childhood health (n=1,929)**	**Poor childhood health (n=176)**	**Difference**	**Statistical sig. of difference**
Self-reported health	56.5%	42.5%	-14.0%	*0.003*
Health compared to peers	80.4%	70.8%	-9.6%	*0.015*
Good perceived health (card)	83.2%	75.1%	-8.1%	*0.028*
Expected longevity (>10 years to live)	44.6%	40.0%	-4.6%	*0.385*
Disability scale	20.51	20.67	0.15	*0.297*
Functional disability (>=5 ADL/IADL)	9.9%	12.8%	2.9%	*0.292*
>=3 chronic diseases	28.0%	29.7%	1.7%	*0.669*
Depressed (% with +10 depression symptoms)	5.2%	13.9%	8.7%	*0.000*
Falsification variables				
Attends religious services > 1 / week	54.1%	46.7%	-7.4%	*0.114*
Sees children > 1 / week	79.4%	71.6%	-7.8%	*0.058*
Urban residence (%)	64.9%	51.5%	-15.1%	*0.003*
Below poverty line (%)	13.4%	18.6%	-15.8%	*0.096*
Below extreme poverty line (%)	6.2%	4.0%	-13.1%	*0.246*

** Analysis takes sampling weights into account*
** Statistical significance tests are two-tailed tests of proportions (categorical variables) and t-tests (continuous variables)*

Table 4: Multivariate Analyses for Senior Life Satisfaction on Childhood Conditions

Main Results	**Marginal Effect**	**Standard Error**
LVDM		
Overall child health	-7.1%	*0.041**
Tuberculosis (%)	-4.1%	*0.118*
Rheumatic fever (%)	-1.0%	*0.089*
Poliomyelitis (%)	N/A	*N/A*
Malaria (%)	-2.0%	*0.033*
Asthma/chronic bronchitis (%)	-6.8%	*0.039**
Sensitivity Analyses		
Pooled cross-sectional, 2004		
Poor overall child health	-14.3%	*0.030****
Change-scores		
	Predicted probabilities of Y associated with poor overall child health	
	Predicted Probability	**Standard Error**
Persistent unhappiness	8.2%	*0.020****
Happiness decrease	3.9%	*0.010****
Happiness increase	2.7%	*0.007****
Persistent happiness	-14.8%	*0.035****

** Outcome modeled as follows: persistent dissatisfaction (< very satisfied at both time points); satisfaction decrease (very satisfied - < very satisfied); satisfaction increase (<very satisfied - very satisfied); persistent satisfaction (very satisfied at both timepoints).*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*
** Due to small sample size (n=5), multivariate regressions were not able to assess the association of happiness with poliomyelitis*

Table 5: Interactions of Childhood Health with other Childhood Conditions

Interactions (cross-sectional, first wave)	**Marginal coefficient**	***Standard error***
*Poor overall child health * Childhood poverty*	-1.4%	*0.079*
*Poor overall child health * Single-parent household*	0.3%	*0.077*
*Poor overall child health * Child No Shoes*	6.4%	*0.084*
*Poor overall child health * Child No Electricity*	10.1%	*0.094*
*Poor overall child health * Child No Bathroom*	-10.0%	*0.077*

** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

** Interactions were run on wave 1*

Table 6: Mediation Tests for Mediators and Falsification Variables

	% of total effect mediated by variable	Standard Error
Mediators		
Subjective health		
Self-reported health (LDVM)	31.4%	*0.009***
Health compared to peers(LDVM)	6.4%	*0.006*
Good perceived health (CS, w1)	11.4%	*0.007*
Expected longevity (expects to live >10 years; CS wave 1)	7.9%	*0.007*
Depressed (% with +10 depression symptoms; LDVM)	26.2%	*0.018*
Objective adult health		
>=3 chronic diseases (LDVM)	3.3%	*0.003*
Functional disability (>=5 ADL/IADL; LDVM)	2.2%	*0.005*
Personality variables		
Locus of control (CS, wave 1)	18.3%	*0.005***
Falsification		
Attending church >=1 week (LDVM)	12.9%	*0.004*
Sees adult children >=1 week (LDVM)	0.5%	*0.002*
Poverty (LDVM)	0.8%	*0.002*
Extreme poverty (LDVM)	0.3%	*0.005*
Urban residence (LDVM)	0.3%	*0.002*

** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*
** Dependent variable: happiness; independent variable: childhood overall health; covariates from happiness model and baseline happiness if LDVM specification*

4. USING HAPPINESS TO EVALUATE THE IMPACT OF GOVERNMENT PROGRAMS

Abstract

Inherent to all governmental programs is the goal of improving the happiness of the public. However, happiness has seldom been used as a program evaluation criterion. Using econometric techniques, this study analyzes the impact of a government program on life satisfaction (i.e. happiness). The program being evaluated is Golden Citizen, a national initiative for the elderly in Costa Rica. Using a nationally representative panel of the elderly in Costa Rica (n=2,827), the analyses suggest that the program did not have a statistically significant effect on life satisfaction, but it did have some positive effects on two more commonly measured indicators -healthcare access and poverty. The article discusses the benefits of evaluating government programs in terms of happiness, and the reasons for a possible lack of a relationship between intermediate outcomes and happiness. By using happiness as an outcome measure, researchers can compare interventions from different sectors.

Introduction

Many people would posit that happiness is their ultimate goal[60]. If, as claimed by US President John Adams, "the happiness of society is the end of all government"[7], an important question is whether government programs can promote happiness. Two opposing schools (summarized by Dutt et al.[7]) emerge on this point. The first postulates that by increasing societal welfare, governments increase happiness, especially amongst vulnerable groups, such as the elderly[7]. The alternative perceives governments as "inefficient and self-serving"; consequently citizens feel alienated and their happiness drops (p. 243). This happens because governmental action encourages "collectivization", thereby curtailing individual freedom, and displacing crucial social structures, such as church and family.

A parallel debate concerns the appropriate instruments to increase happiness. One theory is that fiscal policy can reduce economic inequalities (Veenhoven[7]), non-satisfying consumption (Frank; p. 248), and destructive competition for higher incomes (Griffith)—all of which should increase happiness. Another holds that changes in non-monetary dimensions have more lasting impacts on happiness than changes in income; therefore governmental programs should promote life's non-pecuniary dimensions (i.e. health, social relationships, and others) over the economic ones[153]. Such recommendations are based however on the relationship of happiness to these dimensions; the impact of specific interventions on happiness has been seldom explored.

Researchers overwhelmingly evaluate the impact of policies on intermediate outcomes. This approach, which assumes that happiness will increase as a natural result, underlies the use of economic indicators (i.e. GDP per capita) to evaluate the performance of countries, and of sector-specific outcomes (i.e. infant mortality rate; school enrollment rates) to evaluate governmental programs. However, the assumption that positive changes in people's lives will necessarily lead to happiness may not always hold, and can lead to undesired effects. Indeed, average happiness has remained constant in the US despite steady economic growth over two decades [12]; economic growth can also have substantial externalities, such as environmental degradation and deterioration of community relationships [11]. These factors call for happiness to become an additional metric in policy evaluations: by doing so, evaluations could guide to policies that are more relevant to their ultimate goal, at lower social and environmental costs.

The literature on the elderly illustrates the scant use of happiness as an evaluation criterion. Consider the literature on healthcare programs. One stream has assessed their impact on clinical outcomes, such as survival, physical functioning [154-156], cognitive functioning[155,156], and pain [154]. Within this stream, some have focused on more immediate outcomes, such as falls [157], weight loss [158] and exercise[159]. Mental health outcomes have also been the object of extensive scrutiny, as shown by systematic reviews in the fields of depression [160], bereavement [161], and suicide [162]. A second stream of evaluations focuses instead on the use of healthcare services[163], including visits to emergency departments [164], hospitalizations, nursing homes [165], and patient satisfaction [166]. A third stream explores spending levels [167], and the relationship between spending and outcomes [168,169]. Some of these studies [170,171] use composite outcomes that take quality of life (i.e. Quality Adjusted Life Years) and disability (Disability Adjusted Life Years) into account. All this literature sheds light on whether programs improved certain intermediate outcomes. However; whether these programs increased happiness is generally left unanswered. This is surprising, given that it is likely that people prefer a program that makes them happy than a program that promotes the use of appropriate medical care or its cost-effectiveness.

Public authorities are well aware of the importance of happiness: if governmental programs do not make constituents happier, dissatisfaction will grow, and political capital—their most important asset for a politician—can be lost. Recognizing this, some countries are starting to consider happiness an overt policy goal. Since the 1980s, the Bhutanese constitution mandates the State to "promote those conditions that will enable the pursuit of Gross National Happiness"[61]. Gross National Happiness is measured through an index, which rests upon 33 indicators from 9 domains, one of which is psychological wellbeing[172], a concept akin to happiness. The governments of France and the

United Kingdom have recently taken steps to increase the happiness in addition to the wealth of their citizens [9,61]. As part of those initiatives, the Office of National Statistics of the UK has just published its first national well-being statistics [8]. These actions show that governments are increasingly concerned about the happiness of their citizens.

If public action can affect happiness, then programs that improve happiness (in addition to intermediate outcomes) can be judged more successful than those that only improve intermediate outcomes. Such reasoning provides the rationale for this paper, which provides an example of how happiness can be used to evaluate a government program.

The subject of this evaluation is Golden Citizen (GC), a national program that aims to dignify and improve the quality of life of the Costa Rican elderly[173]. Given its focus on quality of life, the GC program should enhance the happiness of participants. In many ways it is not different from many public programs in the US and other countries. The main benefits it provides—improved access at public institutions, commercial discounts, community activities, exercising—are common objectives of public programs around the world.

Two hypotheses explore these relationships. The first states that the GC program does not affect life satisfaction (i.e. happiness), after adjusting for confounders. The second states that GC does not affect its intermediate outcomes (i.e. improved access to services, reducing poverty status, promoting exercise and facilitating friendship networks), after adjusting for confounders. The intermediate outcomes are process variables, that is, dimensions that should be affected by GC, and that should in turn affect life satisfaction. A rejection of both hypotheses would suggest that GC improved happiness and it did so through its programmatic activities. In such case, this article would provide an example of how a government program can affect happiness. However, if program activities improved intermediate outcomes, but not happiness, then a long held assumption – that a positive impact on intermediate outcomes automatically leads to improved happiness – may be called into question.

The Golden Citizen program

The Golden Citizen program (GC), is open to all residents in Costa Rica aged 65 and older, and has the overall objectives of improving the dignity and quality of life of seniors. GC is a program of the Costa Rican Social Security (CCSS), sole payer and provider of the national health system. The program is coordinated by the Golden Citizen Office, which is part of the CCSS Benefits Division. This central office is supported by five Regional Administrative and Financial Offices, which coordinate a network of 73 social security offices [174]. Individuals can enroll in the program at any of these offices[175].

Members receive a "Golden Card" upon enrollment, which entitles them to program benefits, including increased access to public institutions, commercial discounts, community and recreational activities [173]. By 2007, the GC office had already organized 167 workshops, at which 4,944 workers were trained to assist senior citizens [173]. The program created also 157 agreements with commercial institutions, which were designed to "maintain an adequate infrastructure, preferential seats and other benefits for the elderly"[174]. Discounts to GC members vary between 2%-50%[175], and are available at 6,000 establishments, including pharmacies, supermarkets, theaters, and recreational facilities[174]. Since 2002, GC members enjoy also a "gradual discount in public transportation", whereby they "will travel for free in bus services up to 25 km", will enjoy a 50% discount in distances higher than 26 km, and a 25% discount in distances higher than 50 km"[174]. The GC program conducts also health and culture workshops. The health workshops are conducted by physical education professionals, and raise awareness about the benefits of exercise. The cultural workshops focus on topics such as folkloric dancing, pottery, painting, and flower arrangements [174].

GC has grown over time. Between 2005 and 2007, its membership increased by 25,000 people, and the budget increased five-fold, up to USD 370,000 [173], p. 165

What makes people happy? A short review of the literature

The study of happiness has been always predominant in human thinking. Aristotle, Seneca, Jeremy Bentham, John Stuart Mills and Adam Smith have all published volumes on this topic [12]. However, most of this literature is conceptual; data-driven studies did not emerge until the 1980s. Since then, psychologists, sociologists and economists have published over 3,000 happiness articles [80]. This literature has reached a few points of consensus. First, genetics explains a large share of happiness. Studies on twins have shown it to explain between 30%-80% of happiness across individuals[37]. The impact of genetics occurs through personality traits, such as neuroticism [133],[134], extraversion [133],[134], and locus of control [75], [76]. However, these genetic factors are mostly immutable. Modifiable happiness predictors include health, income, civil status, employment, and social ties [38-40]. Amongst these, health stands [41,42,43,44,] especially among the elderly[45]. Additional important dimensions include age [38,62], religious beliefs [135] and environmental factors, such as safety, urbanization [38] and the political system[69].

Methods

Conceptual framework

This study uses a general happiness model for hypothesis 1 (i.e. relationship between GC membership and life satisfaction). The dependent variable in the model is life satisfaction, the independent variable of interest is GC membership, and a set of controls selected from our literature review are covariates. Following standard practice, relevant controls must be related to both GC program membership (exposure) and life satisfaction (outcome). Though there is a sizable literature on happiness correlates, no published studies have explored the predictors of GC membership. Therefore, a series of analyses (detailed below) were performed to understand which factors were predictive of GC enrollment. Then, the happiness model selected first the main predictors of elderly happiness from the literature, and from these, those related to GC membership. The association of these predictors with GC membership was verified through probit regressions.

The happiness model uses socio-demographic variables (age, sex, relationship status, educational level, urban residence, region of residence, housing assets [38], enrollment in other social programs, pension status, health (metabolic syndrome, chronic diseases, disability status; depression) and personality variables (locus of control [75,76]). . Originally, we had considered other variables typically related to life satisfaction that could be also related to GC membership; however they were not included because either they affect very few respondents or are unrelated to GC membership. These are unemployment,[38,50,72,176] residence in a reliable community—a measure of social capital [177], and two lifestyle variables of relevance in Latin America: how often the respondent sees his or her children [78], and weekly church attendance [79]. Household-level variables had been also considered. These include the number of people in the household, the number of household residents over 65, and the educational level of the spouse. Given that no theoretical foundations support their relationship with GC membership, they were excluded from our estimations.

Happiness and GC membership may be also affected by unobservable characteristics such as personality, culture, and life events. Unless properly adjusted for, these factors become part of the error term, leading to biased and inefficient estimates [111]. Unobservable factors can be time-variant or time-invariant. As discussed in the introduction, time-invariant unobservable factors (i.e. genetics and personality) are crucial in explaining happiness; it is therefore crucial to control for them [80]. We were able to adjust for the influence of these unobservable characteristics through various econometric techniques.

The second hypothesis explores whether GC membership affects process variables. These are healthcare access (which should result from preferential attention), improved income (commercial discounts), exercise [38,51,149] and friendship networks [41,178]. The association of GC membership with each of these will be examined, after adjusting for confounders. In order to test these channels, we have used the exogenous covariates of our happiness model. These are age, sex, educational level, and residence (urban and region of residence).

The data: CRELES

This study uses the "Costa Rican Study on Longevity and Healthy Aging"(CRELES) database [82]. CRELES is a nationally representative panel dataset of individuals 60 years and older in Costa Rica. CRELES covers self-reported physical health, psychological health, living conditions, health behavior, health care utilization, social support, and socioeconomic status. The dataset consists of two rounds of interviews. The first, completed between November 2004 and September 2006, includes 2,827 people. The second, completed between November 2006 and July 2008, contains 2,364 individuals from the first wave. Attrition was therefore 16% (n=463 people). The majority of these (10%) died before the second wave was conducted, and the remainder (6%) could not be located for the follow-up interview[82]. Bivariate analyses revealed that censored and non-censored individuals were not statistically different in terms of life satisfaction, socio-demographic, and health. Hence, attrition-generated bias should be modest.

CRELES is modeled after the Health and Retirement Survey from the United States. Hence, it uses a complex sampling design. To create the survey, a master file of 9,600 individuals born before 1946 was first constructed. This required stratifying all 55+ individuals from the 2000 Census of Population into 5-year age groups, and sampling randomly from these strata, while oversampling the oldest-old (95 years and older). As a consequence of oversampling, sampling fractions varied between 1% (for those born between 1941 and 1945) and 100% (those born before 1905). Then, the master file was divided into the government's already designated 102 Health Areas, 60 of which were selected randomly for the first interview. The resulting sampling frame consisted of almost 5,000 individuals, covering 59% of the national land area. Of these, 2,827 could be located for the first interview, resulting in a non-response rate of 43.3%. Non-responses break down as follows: by the first contact, 19% of interview candidates were already dead, 18% could not be located (due chiefly to inaccurate addresses), 2% had moved, and 4% rejected (directly or indirectly) to be interviewed [81]. The vast majority of the 20% (18%+2%) that could not be located differed from the rest of the sampling frame in terms of age, social class and urban residence. In order to correct for non-response and oversampling bias [83], CRELES incorporates a set of weights. These "allow the replication of the structure for sex, age, residence and education of the whole 2005 population of Costa Rica born in 1945 or before"[81]. Our statistical estimations incorporate these weights as appropriate.

To validate the dataset, the survey team verified the concordance on nine key measures (age, education, etc.) between CRELES and EPHM, another major survey. EHPM (Costa Rican Household Survey for Multiple Purposes) is a nationally representative survey of 12,000 households conducted yearly by the National Statistics and Census Institute of Costa Rica (INEC)[81]. The two surveys had very similar mean values in all 9 variables (<3.5 point difference[82]). A statistically significant difference arose only in the percentage of respondents who are head of household. This difference is however attributed to the fact that in "CRELES the informant is the same older person while in EHPM it can be another person with a different perception on who is the head of the household."[81]. The high coherence between both surveys suggests that CRELES should not suffer from significant response biases.

Another potential bias, typical of elderly populations[84] lies in the use of proxy interviews. At CRELES, 703 interviews in the first wave, and 676 in the second required a proxy due to incapacity to communicate or cognitive impairment of the main respondent[73]. These interviews contain no self-reported variables such as life satisfaction. Given that cognitive impairment is a major reason for proxy interviews, and that assessing life satisfaction requires self-awareness by the interviewee, the exclusion of these individuals seems necessary for analytic purposes. Concerns could be raised about whether excluding proxy respondents could induce bias. This is because cognitive impairment in the elderly has been linked with depression[179], and hence lower life satisfaction. On the other hand, cognitive impairment is unrelated to GC membership in our sample; excluding respondents with high cognitive impairment should therefore not induce bias. Proxy interviews were therefore excluded from our analytic sample, which results in final study populations of 2,111 observations in round 1, and 1,684 observations in round 2.

The variables

The main outcome measure is life satisfaction. Life satisfaction is one of the definitions for happiness [39]. In contrast to other happiness measures, such as positive affect or instant happiness, life satisfaction is more cognitive and hence less vulnerable to transient moods. In fact, stability is the key defining feature of life satisfaction versus other happiness measures[39]. Therefore, of all measurements, life satisfaction may be the most appropriate to assess the impact of a program whose effects can accumulate over time, such as GC. Hence, and in consonance with the majority of the literature, this article uses happiness and life satisfaction interchangeably.

The dependent variable was extracted from the most common question in happiness research [12,89]: "in general, how do you feel about your life? very satisfied, somewhat satisfied, somewhat unsatisfied, very unsatisfied". The same question was used in other large-scale studies[44,49,58,72,86-88]. Similar to other studies, responses presented a ceiling effect—74.7% of respondents in round 1, and 77.8% in round 2 reported feeling very satisfied. Power calculations determined that variability in some categories of the outcome would be insufficient for statistical inference. Thus, this variable was dichotomized into being very satisfied vs. less than very satisfied.

The independent variable of interest is Golden Citizen membership. This variable determines program exposure, and equals 1 if the respondent answered affirmatively to receiving support or services from the Golden Citizen program (=0 otherwise). As established in the conceptual framework, covariates are divided into socio-demographic, health, and personality. Socio-demographic variables include age (categorized into 5 year groups); sex; relationship status (=1 if in a relationship); urban residence (dichotomous); health region code (categorical, 7 health regions); educational level (4 categories: no education, primary, secondary, and tertiary); pension status (dichotomous); enrollment in other social programs (dichotomous) and housing assets. The housing assets index is a proxy for permanent income [102]. This index is a weighted sum [103](range 0–32) of a set of questions on housing quality (quality and materials of floor, ceiling) and home assets (room for cooking, number of telephones, of computers, etc.). In this approach, each asset is assigned a weight, which is the inverse of the number of respondents owning that asset. Hence, asset index=$\sum$asset*w, and a higher score reflects higher asset value.

The health covariates are disability, metabolic syndrome, chronic diseases, and depression. Disability was recoded from 14 items on Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL). The 10 ADL questions asked for difficulties with basic functioning (walking, bathing, eating, going to bed, using toilet, and cutting toenails), and the four IADL questions asked for difficulties with more complicated tasks (preparing food, managing money, shopping, and taking medications). The sum of these items provides a reliable measure [93]; respondents were considered disabled if unable to perform 5 or more activities of daily living. The metabolic syndrome variable (dichotomous) is based on the criteria of the International Diabetes Federation [94]. Individuals were classified with metabolic syndrome if they had abdominal obesity (waist-hip ratio>1) and any two of the following: high triglycerides (>150 mg/dL); low HDL cholesterol (<40 mg/dL in males; < 50 mg/dL in females); hypertension (>130 systolic; > 85 diastolic), or diabetes. The chronic disease variable denotes those with >= 3 chronic diseases as measured by a count of self-reported cancer, heart attack, heart disease, stroke, hypertension, high blood lipids, diabetes, lung disease, arthritis, and osteoporosis. The depression variable (dichotomous) reflects whether an individual is clinically depressed according to the 15-item Yesavage Geriatric Depression Scale (GDS-15). Individuals with

>= 10 depression symptoms were classified as depressed, as established in a psychometric validation in Latin America[91].

The personality variable is locus of control, which is linked to happiness [39]. Locus of control reflects the degree at which respondents feel in control of their lives (i.e. internal control). The locus of control score (continuous) sums 7-items; higher scores would reflect higher internal control [98].

The intermediate outcomes (hypothesis 2) are healthcare access, income/poverty status, exercising, and friendship networks. Measures of healthcare access are whether the respondent waited =<30 min to be seen at the doctor's office (dichotomous), and whether they were visited by a primary care team in the last 12 months (also dichotomous). The second intermediate measure is poverty status. Poverty status reflects current income flows, which are more prone to be influenced by the GC program than permanent income (i.e. housing assets). Calculating incomes entailed estimating personal income, and adjusting it by the number of household members and the cost of housing. First, personal income was estimated separately for the respondent and their spouse. These are the sum of income from work, pension, and transfers from others. Given the absence of income data for other household members, household income was assumed as the sum of the personal incomes of the respondent and their spouse. Household income was then divided by $\sqrt{2}$ if the respondent lived with their spouse[99]. Such adjustment is warranted due to the recognition that two people in the same household consume each marginally less than each person alone. Housing values (i.e. how much the respondent would receive from renting the house they live in and own) were added to incomes. This adjustment is standard in national surveys [180]. Poverty and extreme poverty (both dichotomous) were determined by applying national poverty thresholds for 2004 and 2007[101] to the estimated incomes.

Other intermediate outcomes are exercising and friendship networks. Exercising (dichotomous) reflects an affirmative response to "exercise regularly or do other physically rigorous activities like sports, jogging, dancing, or heavy work, three times a week" over the last 12 months. Finally, a set of variables examine the social relationships of respondents. The first (continuous) is the number of friends reported. Then, to identify social isolation, this variable was also dichotomized into having any friends or not. To identify intensity, the last variable asks whether the respondent talks to their friends at least once a week (dichotomous). Contrary to the rest of variables, the friendship variables are asked only in the second wave of CRELES, which limits the choice of econometric methods for their estimation.

Analytic approach

To understand the data, the bivariate relationships in the analytic sample were reviewed. The first step was to understand the differences between GC members and nonmembers, particularly in regard to life satisfaction. This proceeded by exploring the demographic characteristics at both interview periods (table 1). Then, a comparison took place of the baseline distribution of the main life satisfaction predictors in GC members vs. non- members (table 2). Then, it was necessary to understand the bivariate relationships along the causal chain: GC membership, intermediate outcomes, and final outcome (life satisfaction). These are explored in table 3 (relationship between GC membership and intermediate outcomes) and table 4 (relationship between intermediate outcomes and life satisfaction). Afterwards, to assess the extent of program take-up, a comparison of GC membership took place between both waves (table 5). Finally, we explored the predictors of GC membership using probit regressions. Originally, we had tried the bivariate probit models with partial observability[181] to this end. However, these failed to converge in every case (a common problem in bivariate probits). For this reason, we opted for simple probits. As we had no theoretical model of which happiness predictors should be associated with GC membership, we added variables to our program take-up models in a block wise manner. First, we defined a set of core variables. These are socio-economic: age, sex, educational level, housing assets, place of residence (urban residence and area of residence), and relationship status. Pension status was also included as a core variable, because it is a strong predictor of GC membership, and is also a guaranteed source of income, which should affect wellbeing. Then, using log-rank tests, we tested the statistical significance of enrollment in other governmental programs; health (disability, chronic diseases, metabolic syndrome; depression) personality (locus of control), social life dimensions (frequency of seeing children; frequency of religious attendance), and household dimensions (number of people over 65, educational level of spouse, number of people in household). Amongst these, only programs, health, and personality variables were significant. Therefore, the final happiness model used for hypothesis 1 includes these three significant dimensions as well as the core happiness predictors.

There were several concerns in choosing the econometric approach[111,182]. Four key dimensions, listed in terms of importance, were taken into account. The first is endogeneity, a major threat to causality in observational studies[111], because it leads to biased and inefficient estimates, rendering them "causally uninterpretable"[111]. Typically, endogeneity results from simultaneous causation between exposure and outcome, and from the influence of unobservable confounders[111]. These two sources of endogeneity are crucial when individuals can self-select into the program being evaluated, such as GC. Removing endogeneity is therefore a primary concern. Second, within the second source of endogeneity (i.e. influence of unobservable confounders), the most important in happiness research are time-invariant unobservable factors, as they include genetics and personality[80]. Third, changes in life satisfaction should be depicted as accurately as possible. That is, we consider individuals who were satisfied at both timepoints to be substantially different from those who were dissatisfied; these groups should be therefore analyzed separately. Fourth, for causality to occur, cause must precede outcome [111].

Various methods, each with its own strengths and weaknesses, can be used to address the previous methodological dimensions. Instrumental variable (IV) estimation and regression discontinuity (RDD) are commonly used to treat endogeneity. However, these techniques identify only a local average treatment effect, i.e. the effect on a small subset of observations. Therefore, IV and RDD raise generalizability and precision concerns. The first arise because the effect applies to fewer observations; it is difficult to generalize findings to all individuals. Precision is a concern because as IV and RDD use fewer observations, their standard errors are artificially wide. In addition, these techniques rely on assumptions that cannot be always fulfilled. Additional models to be considered include fixed-effects, differences in differences (DiD), lagged dependent variables models (LDVM), change-scores, and propensity scores. Fixed-effects models can accurately remove the effect of time-invariants (second dimension). However, they are known to be biased in their ordered probit form[80], which is the specification necessary for this study, as the outcome is categorical. DiD methods can cancel out the effect of time-invariant factors (dimension 2) and fixed effects common to all individuals, but cannot model the outcome as desired (dimension 3) or ensure that cause precedes outcome (dimension 4). Change-scores and LDVM models do ensure that cause precedes outcome; however adjust only imperfectly for unobservable factors [106], particularly time-invariant [110]. Within these two, change-scores have the advantage of modeling separately those who stay satisfied vs. dissatisfied at both time points, thereby complying with the third dimension. Propensity scores can be used to enhance covariate balance—and hence comparability—between both groups. However, they also assume that there is no unobservable confounding– an unrealistic assumption outside of a randomized trial. Propensity scores are therefore an imperfect technique to remove endogeneity (dimension 1). Hence, all models have their pros and cons, and no approach is methodologically ideal.

To balance out these trade-offs, we selected IV as primary and RDD as secondary analyses. Under the right assumptions, both methods can remove endogeneity, the foremost methodological concern. Sensitivity analyses are DiD, LDVM, and change scores in their unadjusted and propensity-score adjusted forms. Amongst these, DiD have the key advantage of removing time-invariant fixed-effects, which are key in happiness research. Consistency across specifications would lend validity to our results.

Statistical analyses for hypothesis 1 (effect of GC on happiness)

The main analysis is instrumental variable estimation (IV). The instrument is an exogenous variable that affects the outcome only through the exposure of interest. As discussed, the exposure variable (i.e. GC membership) is endogenous, and therefore cannot be estimated with standard techniques. IV estimation uses the instrument to model first the exogenous exposure (first-stage equation); and then estimates the effect of this modeled exposure on the outcome (reduced-form equation). By identifying the effect only for those individuals who joined GC as predicted by the (exogenous) instrument, estimates are consistent, and hence free of endogeneity. For this to occur, the instrument must fulfill two key conditions: a) strength, which implies that it is sufficiently correlated with the endogenous explanatory variable, conditional on the other variables; and b) validity, whereby the instrument is uncorrelated with the error term in the reduced-form equation. Under such assumption, the instrument would be uncorrelated with both the outcome and observed factors, i.e. endogeneity would be removed.

The instrument in our IV estimation is the percentage of neighbors enrolled in the program. The instrument was constructed by dividing the number of GC enrollees over total respondents within each health area (total of 60 health areas). Conceptually, this should be a sound instrument, as high GC-density areas should not be systematically different from the rest, and individuals in those areas should feel more compelled to join the program than those living in areas with low GC penetration. Standard statistical tests were used to verify the instrument´s strength (Crag-McDonald F-test) and validity (Wu-Hausman).

The IV estimations use the exogenous covariates from the happiness model. These include age, sex, educational level, and residence (urban and region). The models were estimated using bivariate probit (-biprobit- command in Stata 11), the specification indicated when exposure and outcome are both binary [183,184]. However, results must be taken with caution, as bivariate probits tend to result in artificially narrow standard errors.

Accordingly, IV models used the following form:

Probit (Very Satisfied)$=\beta_0+\beta_1(GC^*)+\beta_2(Age)+\beta_3(Sex)+\beta_4(\text{Years of Education})+\beta_5(\text{Urban Residence})+\beta_6(\text{Region of Residence})+\varepsilon$ (reduced-form equation);

Probit $(GC^*)=\beta_0+\beta_1(Z)+\beta_2(Age)+\beta_3(Sex)+\beta_4(\text{Years of Education})+\beta_5(\text{Urban Residence})+\beta_6(\text{Region of Residence})+\varepsilon$ (first-stage equation);

where Z=% of neighbors currently enrolled in the GC program and GC=instrumented program participation.*

IV estimations were conducted on a pooled cross-sectional database; the outcome is therefore binary (very satisfied vs. less than satisfied). Hence, results are expressed in average marginal effects, which reflect the average probability change in the outcome that results from a one unit increase in the independent variable.

The secondary analysis is non-parametric regression discontinuity (RDD). In RDD, effect identification relies on an exogenous cutoff that drives program participation. By including this cutoff, the error term in the estimation equation contains "no information that might correlate with the grouping variable"[111]. Therefore, the influence of unobservable confounders--i.e. endogeneity—is removed.

Individuals must be at least 65 years old to join the Golden Citizen program; RDD uses this cutoff to model program participation. The effect is identified by comparing local regressions of observations that are just before and just after the cutoff. Hence, a local average treatment effect on compliers is identified [182] per the following equation:

_frd= E[Yi(1) – Yi(0)|unit is a complier and Xi = c].

Given the low number of observations used in the local regressions, RDD suffers from the generalizability and precision concerns mentioned earlier. We tried to address these by pooling both CRELES waves (as if they were a single cross-section) to increase the sample size for the RDD models. Though this may not improve generalizability, the higher sample size should improve the precision of the estimates.

The RDD model has some additional features. First, it is non-parametric, thus making fewer assumptions about the error terms than parametric specifications [185]. Second, it uses a fuzzy design. In contrast to sharp RD, which applies only when all participants after the cutoff are enrolled in the program (i.e. Medicare program in the US; everybody over 65 is enrolled), in fuzzy RDD "the probability of receiving the treatment need not change from zero to one at the threshold"[182]. Indeed, fuzzy RD requires only a "discontinuity in the probability of assignment to the treatment at the threshold". Therefore, fuzzy RDD is advisable only when a subset of individuals sign up onto the program– a realistic assumption in the case of GC. Third, the RDD model uses a local linear regression kernel. This is advised when "we are interested in the behavior of the regression functions around a single value of the covariate" [182], i.e. in the change in happiness between just before and after age 65. Fourth, our RDD estimations used a specialized command in Stata–rd- 11 that optimizes the estimation bandwidths, i.e. minimizes the squared bias plus variance of the local regressions[112]. This was adequate to ensure a wise balance between the bias and precision of estimates. Fifth, the analyses controlled for survey wave and used robust standard errors. Controlling for survey round was necessary to control for time effects – RDD proceeded on a pooled sample. Robust standard errors were necessary to control for heteroskedasticity.

Finally, several checks were conducted to ensure that the RDD analysis meets its basic assumptions. This entailed first verifying that there is indeed a discontinuity in life satisfaction (outcome) at age 65 (cutoff). Second, we inspected the existence of jumps in life satisfaction at other ages. A discontinuity in life satisfaction at age 65, but not in other ages, would be expected. Third, a jump in GC enrollment (exposure) at age 65 was verified. Fourth, we contrasted our results using various estimation bandwidths. The bandwidths determine the number of observations used in each local regression. As they use fewer observations, narrower bandwidths are associated with lower variances but higher bias[186]. Similar results across various bandwidths would strengthen results. Given RDD estimations are pooled cross-sectional, the outcome was also modeled as being very satisfied vs. less than very satisfied.

Sensitivity analyses

To our knowledge, life satisfaction (i.e. happiness) has not been used before in program evaluation. There is therefore uncertainty about the sensitivity of this variable to government programs. This requires extra care to ensure the robustness of findings. To that end, sensitivity analyses were used. These are difference-in-differences (DiD), lagged-dependent variable models (LDVM), and change-scores—with and without propensity scores.

DiD compares the pre-post differences in the conditional probability of being very satisfied between those who joined vs. those who did not join the program. This requires first calculating the pre-post difference within each group, and then comparing these differences between both groups. Under the right circumstances, DiD controls for the effect of time-invariant confounders. However, DiD does rest on two key assumptions. The first is that GC membership is uncorrelated with time-variant confounders. Given that people may self-select into the program, this assumption does not seem fully supported; the results must be therefore interpreted with caution. The second assumption states that the differences between the intervention and control groups stay constant [187]. As the survey timepoints are separated by only two years, this can be reasonably assumed. In order to fortify the findings, propensity scores will be used to enhance covariate balance—and hence reduce confounding—in the analysis. DiD estimations were also conducted on a pooled sample, but restricted to those not in GC at baseline. This restriction is necessary as DiD identifies the effect of joining the program; the analytic sample must refer only to those not in GC at baseline. The general design of DiD is:

$$Probit(Very\ satisfied) = \beta + \beta(GC) + \beta(t) + \beta(GC*t) + \beta(X) + \varepsilon$$

In this specification GC=1 if the individual joined the program between both waves (and 0 if not), and t=CRELES round. The interaction term (GC*t) is the

DiD estimator, which reflects the effect on life satisfaction of joining the GC program. We calculated the interaction term with the Stata –inteff- command [188], which computes interaction coefficients for nonlinear models. In the equation, X is a vector of the happiness model covariates available in both waves. These are age, sex, civil status, education level, urban residence, area of residence, pension status, assets, disability, chronic disease, metabolic syndrome, and depression. Finally, and following standard practice, the DiD models controlled for the autocorrelations of the error term.

The next sensitivity analysis is LDVM. These models regress second wave life satisfaction on second wave predictors and baseline life satisfaction (Y_2 on X_1 and Y_1). By including baseline life satisfaction, LDVM adjust for reverse causality and some unobservable confounders [106]; however they offer no guarantee of canceling out the effect of all time-invariant unobservables, such as genetics. LDVM are considered as an alternative to change-scores [110], and, in our specifications, conceive the outcome as dichotomous (very satisfied vs. less than very satisfied). Hence, the LDVM models adopt the following form:

$$Probit(Very\ satisfied_2) = \beta_0 + \beta_1(GC_2) + \beta_2(X_2+) + \beta_3\ (Very\ satisfied_1) + \varepsilon$$

Where the subscript denotes the survey round, and the X vector refers to the same covariates as in the difference-in-differences model. As the outcome is dichotomous, results are expressed as average marginal effects.

The third sensitivity analysis is change-scores. This model measures how changes in life satisfaction relate to baseline covariates (Y_2-Y_1 on X_1), including GC status. Hence, this specification ensures that cause precedes outcome, and also adjusts for the effect of historical unobservable confounders [106]. In the change-scores models, life satisfaction changes were ordered from the least to the most desirable. They were therefore coded 0 for those who were less than very satisfied at both time points, 1 if life satisfaction decreased (Very Satisfied – Less Very Satisfied), 2 if life satisfaction increased (Less Very Satisfied – Very Satisfied), and 3 if individuals were very satisfied at both waves. Changes in the outcome were modeled in the same manner elsewhere [107]. As the outcome categories follow a natural order, these models adopted an ordered probit specification:

$$\textit{Probit}\ (\Delta \textit{Life Satisfaction}) = \beta + \beta(\textit{Golden Citizen}_1) + \beta(X_1) + \varepsilon$$

Where the subscript denotes the survey wave, and the X vector includes the covariates available in the first wave. These are the same as in the DiD and LDVM, with two additional variables: programs and locus of control. Given that the outcome is 4-level categorical, the results are expressed in terms of predicted probabilities, using satisfied-satisfied as the reference group. These provide the probability change in a given level of the outcome, resulting from a one unit-increase in the independent variable. Therefore, results express the probability of being in a certain category (vs. satisfied-satisfied) that results from GC membership.

As mentioned, all sensitivity analyses (Change-Scores, LVDM, DiD) were re-estimated with propensity scores. Propensity scores express the probability of receiving the treatment, given a set of predictors. Hence, the development of PS relies on a logit model:

$e_i = P(Ti = 1|Xi)$; where e_i is the propensity score

Propensity scores create intervention and control groups that are very similar in the probability of receiving treatment as predicted by observable characteristics [189]. Propensity scores have two main features. First, they facilitate group comparison -- the distribution of observed variables becomes more similar between both groups. Second, propensity scores assume that there is no unmeasured confounding[189]. If this assumption is fulfilled--which cannot be taken for granted in observational studies —propensity scores create intervention and control groups similar to those of a randomized trial. In such case (which is very unlikely), propensity scores would remove endogeneity.

Originally, two distinct sets of propensity scores were created. These are simple and full model propensity scores. Simple model propensity scores are based only on exogenous variables thought to affect both GC membership and life satisfaction: age (categorical, 5 year groups), sex, educational level, urban residence, and area of residence. Wave 2 propensity scores include these variables + baseline GC membership. Full model propensity scores included both exogenous and endogenous predictors. These are the ones used in the simple models as well as: enrollment status in other government programs, poverty status, pension status, disability, having four chronic diseases, depression, and locus of control. As the results across estimations (change-scores, LDVM, DiD) were very similar in both, we report only simple model PS. The results for full model propensity scores are available upon request.

This study uses propensity scores in two different ways. These are nearest neighbor matching (1:k PSM) and weighting by odds, both of which identify the average treatment effect on the treated (ATT). The first approach (1:k PSM) is used in the change-scores and LDVMs analyses, and creates intervention and control groups where each observation in one group (i.e. intervention) is matched with another in the other group (i.e. control) that has a similar propensity score. Hence, estimations compare very similar groups as long as no unmeasured confounding exists. The second approach (weights by odds) is used in the differences-in-differences analysis. Here individual observations are not matched; it is the covariate distribution across the whole group that matters. Indeed, the weights by odds make the distribution of covariates across the control group similar to the treatment group [190]. In this estimation, treatment observations receive a weight w=1, and controls w=(ei /1-ei), where e_i is the propensity score. Compared to PSM, weighting by odds has the advantage of not excluding any observations, therefore using a larger sample.

The propensity scores analyses involved the following steps:

- Restriction of analytic sample to individuals >=65 years old, to enhance group comparability.
- Estimation of the propensity scores.
- Development of the PS-based technique to enhance the comparability (i.e. covariate distribution) between groups.
- Nearest neighbor 1:k matching was executed with replacement and the common support restriction. Matching with replacement improves the covariate balance (though it may increase the variance). The common support restriction entails dropping propensity score outliers, i.e. "treatment observations whose pscore [propensity score]is higher than the maximum or less than the minimum pscore of the controls"[191]. This condition may improve the identification of the (ATT) effect[192], as it ensures a better overlap between the propensity scores of both groups. Weights by odds were developed manually using the formula above.
- Balance check of the propensity scores. Particular attention was given to overall and within variable standardized bias. A standardized bias of <=5% was considered adequate. We also verified the descriptive statistics of the propensity scores and histograms to verify overlap of PS between groups. For the weights by odds analyses, we checked the distribution of the weights through boxplots and histograms. Outliers were trimmed at the 95% level.
- Estimation of the study designs (change-scores, LDVM, and DiD), with and without the propensity scores.

Estimation of the intermediate outcomes (hypothesis 2)

We used the same study designs to assess the impact of GC on intermediate outcomes. These are healthcare access (waiting for <=30 min at doctor office; being visited by a primary care team in the last 12 months), poverty and extreme poverty status, exercising (exercises >= 3 days/ week; y/n), and social relationships (having any friends vs. not; number of friends; speaks to friends at least once a week). Hence, the analyses for the intermediate outcomes include IV, RDD, DiD, LDVM and change-scores.

The RDD models for the intermediate outcomes adopt the same functional form as for the primary analysis (i.e. no covariates). The rest (IV, DiD, LDVM and change-scores) used only the exogenous predictors from the happiness model. These are age, sex, educational level, urban residence and area of residence.

Data imputation

At CRELES, missing data is not a significant problem: only 5% of observations were missing two or more happiness predictors. Variables with missing data were imputed. Three variables (depression scale, locus of control, and allostatic index) were missing in >5% of observations. Imputation proceeded through two different methods. First, single imputation proceeded for index variables (i.e. variables that are the sum of subvariables). These are depression, disability, locus of control, and assets. Single imputation was appropriate as information was available on the majority of subvariables in each index (only <5% of observations were missing >20% of items in the index). The rest (chronic diseases, allostatic index, frequency of seeing children, metabolic syndrome, and golden citizen membership) were imputed through multiple imputation chained equations (MICE)[104]. MICE predicts missing values based on other variables with non-missing values; it imputes first the variable with the least number of missing observations; then the second, and so on. All variables are cycled through in such manner, and the whole process is repeated (usually 5-10 times)until a dataset with converged estimates has been reached. This process produces a single imputed dataset with no missing data[105].

Additional measures enhance the quality of the statistical models. Robust standard errors address potential heteroskedasticity, and variance inflation factors were checked to ensure the absence of collinearity.

Results

Table 1 reflects the changes in our population between both waves. Significant changes occurred in life satisfaction, age, household composition and income. The proportion of those who feel very satisfied increased by over 3 percentage points ($p<0.05$). This is in accordance with existing literature, which finds that after a mid-age trough, individuals usually get usually happier as they age[38,39]. The average age increased by 1.16 years, reflecting overall aging of the sample and attrition at older ages ($p<0.001$). There was also a small reduction in average household inhabitants (0.17; $p<0.05$), most likely a consequence of widowing. Finally, the increases in personal income likely reflect a governmental raise of pensions for the poor. Indeed, between 2006 and July 2007 the Arias administration raised the lowest pensions from 16,000 colones (US$32) to 50,000 (US $100)[73]. These income changes are mirrored by a (statistically nonsigificant) 1.3% reduction in the proportion of poor in the sample. The rest of changes are of low magnitude, insignificant, and can also be explained by the aging process (i.e. increase in proportion of females; decrease in people in a relationship, etc.). Hence, it seems that other than the governmental pension increases, most changes in our sample are related to aging.

Table 2 shows the differences between GC members and non-members at baseline. The numbers suggest that differences may be primarily due to age. First, individuals must be 65 to join GC, which explains the higher ages in the GC group (5.41 difference, $p<0.001$). The age differential also could explain the higher proportion of widowed, disabled and pensioned individuals in the GC group. Indeed, all these dimensions have significant relationships with age ($p<0.001$). Pension status may be the main catalyst for GC membership: as GC is a social security program, it is reasonable that pensioned individuals are more encouraged to join, explaining the 25 point difference in pension status between both groups ($p<0.001$). Hence, age seems to be the main driver of the differences between GC members and non-members, including the 25% difference in pensioned individuals between both groups.

We then explored whether the intermediate outcomes are related with GC membership and life satisfaction. Tables 3 and 4 display such relationships respectively. Table 3 shows the relationships of the intermediate outcomes with GC membership. The table shows that GC members wait less at the doctor's office, are less likely to live in extreme poverty, and exercise less (all $p<0.05$). Initially, we would have expected Golden Citizen members to be the ones who exercise more; however, they are also older, which may explain their lower levels of exercise. Frequency of visits by primary care teams, average income, moderate poverty, and social relationships are statistically unrelated to GC membership. Table 4 displays the relationships of the intermediate outcomes with life satisfaction. These numbers show that except for primary care visits, all intermediate outcomes have bivariate associations with life satisfaction ($p<0.05$). Such relationships suggest that if GC impacts its intermediate outcomes, it should be also impacting life satisfaction. Subsequent multivariate analyses will explore these associations in more detail.

The propensity scores showed desirable properties. First, as reflected in table 5, matching decreased considerably the standardized biases. Mean standardized bias is an aggregate measure of group overlap on a set of covariates; the lower the bias, the higher the overlap. It is therefore a standard diagnostic for propensity scores. Formally, it expresses the mean difference between both groups divided by the square root of the mean sample variance of both groups[193]. A standardized bias of < 5% is desirable for inference purposes (i.e. reflects good overlap). As table 5 shows, the mean standardized bias (i.e. for all variables) fell from 7.1 to 0.2 after matching at baseline. Wave 2 showed also substantial reductions, from a mean standardized bias of 17.5 to 4.8. All these measures fall below the 5% standard cutoff.

The standardized bias figures for individual variables (available upon request) show also desirable properties. At baseline, these are all <1% after matching, and none of the observations were off common support, i.e. no treatment observations had a propensity score above the maximum or below the minimum of the propensity score of the controls. Such figures change slightly for wave 2: though the rest of variables have standardized biases < 5%, the standardized bias for urban residence is 9.0% and for area of residence is 9.3%. However, as these variables are not strongly related to the outcome, they should not bias our estimates. In addition, only 10 observations were off common support at round 2. The low number of unmatched observations reflects that PSM resulted in the loss of a residual amount of information. The graphical analyses support the previous notions. Indeed, graphs 1 and 2 demonstrate a good propensity score overlap at both waves between the intervention and comparison groups. The graphs reflect also that propensity scores are considerably higher in the second wave. This is because wave 2 propensity scores include baseline GC status, which is strongly predictive of GC status in the second wave (β=31%; $p<0.0001$). As mentioned in the methods section, further diagnostics explored the distributions of the propensity score weights used in the DID analyses. Exploratory analyses showed that the original distribution of these weights is right tailed, with a median=0.70, mean=0.92 and standard deviation=0.58. Boxplots showed that 125 observations (7.6% of the sample) would be outliers under such distribution. Large weights can result in inefficiency and bias [194]; we thus decided to trim at the 95% level (weight>2.11). This should be acceptable, given that graphs 1 and 2 show that there are few propensity score outliers – eliminating these observations should not substantially bias our estimates. After trimming, boxplots did not reveal any outliers.

Program enrollment

As table 6 reflects, program take-up was considerable: over a quarter of the interviewees joined the program between both waves. Such high enrollment was concurrent to a five-fold GC budget increase between 2005 and 2007. Still, 68.1% of the respondents did not change their Golden Citizen status, where the proportion of those who did not stay enrolled (51.0%) is much higher than those who stay in the program at both rounds (17.1%).

Hypothesis 1: Effect of Golden Citizen on Life Satisfaction

The data do not support that the GC program impacted life satisfaction. Table 7 displays a summary of the findings for hypothesis 1. IV and LDVM suggest a negative relationship between GC membership and life satisfaction. However, RDD and DiD do not find a significant association. In change-scores models, the only models that examine changes over time, baseline GC membership is statistically associated with persistent dissatisfaction, and persistent satisfaction. However, baseline GC membership is also associated with increasing satisfaction. Therefore, the results are inconclusive.

Table 7 summarizes the relationships between GC and life satisfaction. The table shows that according to the IV estimation, GC membership reduces the probability of feeling very satisfied by 4.23% ($p < 0.05$). This analysis seems robust, as the instrument used (% of neighbors enrolled in the GC program) appears strong and valid. The Cragg-Donald Wald F-statistic ($F=89.7$) shows the instrument is strong, i.e. correlated with GC enrollment after adjusting for exogenous confounders ($F>10$ is indicative of a strong instrument). The Wu-

Hausman statistic for validity is not significant (p=0.80), which suggests that the instrument is uncorrelated with the error term, and hence exogenous. Hence, the instrument is adequate for analytic purposes.

However, the secondary analysis (RDD) does not find an effect. Indeed, the RDD results reveal a negative but nonsignificant effect of GC on life satisfaction. These results are to be considered within the validity of the RDD design. First, the data show that after age 65, 50.1% of the observations are enrolled in GC. Given that only a subset of the potential enrollees chose to join the program, estimations must rely on fuzzy RDD. Second, as graph 3 shows there is a jump in GC enrollment at age 65 and no other such discontinuities at other time points. Though there seems to be another discontinuity in program take-up around age 100, this is due only to 7 individuals, and therefore not relevant for analytical purposes. Therefore, the assumption on program enrollment is fulfilled. Third, there is also a jump in the probability of feeling very satisfied at age 65 (graph 4). Although this may be the largest discontinuity in the graph, other such discontinuities occur at ages 62 and 76. We first hypothesized that retirement may explain these three discontinuities (at ages 62, 65 and 76). If this were the case, the validity of the RDD model could be questioned. Unfortunately, CRELES does not provide age of retirement, so we were not able to assess the cause of these discontinuities. However, several facts challenge such notion. First, only a minority (36.9% of seniors nationwide) receive a pension[173]; chart 4.8. Second, in the main pension regimes, individuals may not retire before age 65[173], p.20. In such case, retirement is very unlikely to explain the happiness discontinuity at age 62. Third and most importantly, the effect of retirement on wellbeing (including life satisfaction) is still disputed, with evidence for positive, negative, and neutral effects[195]. Therefore, it is not readily apparent why these other discontinuities occur; it is possible for some of thee to be statistical artifacts. Finally, table 7 shows that the RDD results are robust to various bandwidths. The default is the bandwidth selected by the –rd- command (bandwith=1.7), which minimizes the MSE[196]. Following standard practice, we also report results for half and twice that bandwidth size. The table shows that the effect of GC membership is negative but non-significant in all cases. Hence, a causal effect is not readily apparent in the RDD design.

The sensitivity analyses find also a small negative or no effect. DiD, which (as it in principle cancels out the effect of time-invariant factors) is the most methodologically robust, reveals also non-significant results, whether it be unadjusted ($\beta=0.41$; $p>0.05$) or propensity-score adjusted ($\beta=2.54$; $p>0.05$). However, it is important to keep in mind that DiD models restrict the sample to individuals not in GC at baseline; their study population is therefore slightly different from the rest of estimations. LDVM, in turn, finds a significant negative association when unadjusted ($\beta=-5.40$; $p<0.05$), but becomes nonsignificant once propensity scores are used ($\beta=1.06$; $p>0.1$). The change-scores models (also in table 7) provide a mixed picture as well. The unadjusted change-score models finds that GC members have a higher probability of persistent dissatisfaction (6.07%; $p<0.05$), and a lower probability of persistent satisfaction (-11.59%; $p<0.001$). However, the same table also shows that GC membership is also associated with an increased probability of transitioning into satisfaction (5.15%; $p<0.05$). These patterns repeat themselves in the propensity-score adjusted models. Taken together, all these models (primary, secondary and sensitivity) suggest that GC enrollment is unrelated or may have a small negative effect on life satisfaction.

Hypothesis 2: Effect of Golden Citizen on Intermediate Outcomes

Most evaluations focus on intermediate outcomes. In this section we focus on whether the GC program impacts these measures. The analysis concentrates on four intermediate outcomes: - healthcare access, poverty, exercising, and friendship networks. With the exception of RDD, all models (IV, DiD, LDVM and change-scores) find impacts on poverty. In addition, IV and change-scores find impacts on healthcare access.

Table 8 summarizes the impact of GC on intermediate outcomes. IV and change scores analyses support that GC improves healthcare access, but only in its first variant – probability of waiting < 30 min at the doctor’s office. Further, these impacts are of low magnitude. In the IV model, GC membership improves such probability by 2.0% ($p<0.05$). In the change-scores models, baseline GC membership is associated with a 3.91% higher chance of persistent adequate healthcare access (i.e. waiting < 30 min in both time periods), and lower probability of persistent inadequate access ($\beta=-4.9\%$; $p<0.05$). Notably, GC membership showed no association whatsoever to the other healthcare access variable – being visited by primary care teams. The impacts on poverty seem more robust. In the IV analyses, GC membership reduces the probability of poverty by 8.3% ($p<0.05$), and extreme poverty by 5.6% ($p<0.05$). The DiD and LDVM models support effects of similar magnitude. Taken as a whole, these sensitivity analyses suggest that joining GC is associated with a 5.9% average decrease in the probability of poverty ($p<0.05$), and a 3.8% reduction in extreme poverty. In addition, the change-scores models suggest that GC may be mainly associated with keeping people out, rather than transitioning into or out of poverty. Indeed, GC enrollment is associated with 7.0% ($p<0.05$) lower probability of persistent poverty (7.0; $p<0.05$). GC is also associated with a lower probability of transitioning into and out of poverty (in the -3.3% -1.8%, range; see table). At first, it seems contradictory that opposite transitions should be associated with GC. However the estimates are of low magnitude, so the real effect on poverty transitions is likely to be small. Hence, GC seems mainly to protect from falling into poverty—rather than transitioning into or out of it. Hence, the evidence supports an effect of GC on poverty, and potentially on healthcare access.

The two other intermediate outcomes are exercising and friendship networks. In table 8, only LDVM finds an increased probability of exercising for GC members (β=6%; $p<0.05$). As this variable is not significant in the rest of the models, it cannot be taken as conclusive. Impacts on friendship networks are found, but only intermittently. The IV model finds that GC is associated with higher intensity of contact with friends (3.9% increased probability of seeing friends at least once per week; $p<0.05$). The LDVM model shows that GC is associated with a 1.4 point increase in the number of friends ($p<0.05$). The change-scores shows that the program is associated with an increased probability of having at least one friend; however the marginal probability is low, and borderline significant (β =1.5%; $p<0.1$). In consequence, different models have found associations of GC with different dimensions of friendship—the results are not fully consistent. Therefore, it cannot be concluded that GC had a significant impact on exercising or friendship networks.

Finally, the same table shows that none of the RDD models were statistically significant. This may be explained by the fact that none of these variables displayed a jump at the 65 year old cutoff, which decreases the probability of identifying an effect.

As a summary, the analyses on the intermediate outcomes suggest that GC reduces poverty and may marginally improve health care access, but has no discernible effects on exercising and social relationships.

Discussion

The Golden Citizen program was not successful in improving life satisfaction (i.e. happiness). In fact, under some model specifications it has a negative association with life satisfaction. However, it appears to have had a positive impact poverty and –to a lesser degree-- healthcare access. These are examples of the intermediate indicators that are traditionally used to evaluate public programs. Other important process variables, such as home visits by primary care teams, exercising or extending friendship networks bear little connection or are completely unrelated to program participation. Therefore, though GC has impacted some important intermediary outcomes, it did not impact happiness.

An evaluation focusing on intermediate outcomes would have deemed the program successful. This would be in consonance with the literature discussed earlier, which often finds that government programs affect intermediate outcomes. However, if the program is evaluated in terms of happiness (i.e. its ultimate outcome) then GC was not successful. Hence, the main finding of this paper is that by only evaluating intermediate outcomes, researchers may conclude that programs are successful though there may not be an impact on the ultimate outcome – happiness.

The important question is therefore, how can a program improve the intermediate indicators, while having no impact (or even a negative impact) on happiness? This contradictory finding may be first explained by strong selection bias: GC members could be different from non-members in unobservable characteristics that render them less satisfied. For example, pessimistic individuals may tend to join the program, or the program may have higher penetration in depressed neighborhoods. However, the econometric techniques used in this study (particularly IV and DiD) should reasonably adjust for such biases. A second explanation may lie in the fact that the Golden Citizen program has its strongest effects on income. There is substantial literature that people´s happiness adapts to income over time[20,197,198,18]; by acting on income, the Golden Citizen program may initially affect happiness, but such effect would dissipate as people adapt. Adaptation would be more likely to occur when the effects on income are not large, as in the present case. In addition, the evidence suggests that the GC program was more effective at keeping people out of poverty, than in moving them out of poverty. Had the effects been larger, or had DC been more successful in moving people out of poverty, then the effects on happiness may have been stronger.

Another potential explanation is that the intermediate outcome measures may not be correlated with life satisfaction. In such case, these dimensions would not bridge the effect of GC on life satisfaction. However the robust bivariate relationships between the intermediate outcomes and happiness (table 3) reject such alternative. Another possibility lies on the potential existence of negative program externalities not captured by our data. Though GC helps alleviate poverty (and marginally healthcare access), it is possible that enjoying the program benefits requires cumbersome arrangements, such as lengthy enrollment processes or administrative challenges. Field research could shed light on whether such implementation aspects could create an unfavorable final balance (in terms of life satisfaction) for recipients. A final possibility is that the relationship between the intermediate variables and life satisfaction could be mediated by other factors. Indeed, researchers have found important mediators between objective circumstances and quality of life (of which happiness is a subcomponent), including "positive cognitive bias, homeostasis, unrealistic optimism, positive illusions and illusion of control"[199]. Were these strong mediators, then they could impair the effect of the intermediate variables on life satisfaction . In this sense, life satisfaction and the intermediate outcomes would act as "separate thermometers measuring different processes in the same body, but in related ways"[199].

Can governmental policy affect people´s happiness? Setpoint theory, which assumes that individuals have a pre-set happiness level to which they revert inevitably, provides some guidance [22]. This theory postulates that life events (marriage, changes in income, loss of a loved one, etc.) may temporarily affect happiness levels, but individuals will eventually revert to their fixed baseline [200]. This is because happiness is mainly determined by genetics [37]. If this were true, objective conditions (that is, those external to the individual), which are more prone to government action, would contribute little to happiness. While this seems true for the pecuniary realm (Easterlin 2003) –individuals tend to adapt to income—accumulating evidence supports that people do not adjust as well to non-pecuniary dimensions(losing a loved one, becoming disabled, etc.)[153,201]. Therefore, if government programs had a stronger effect on the latter (i.e. if the Golden Citizen had had stronger effects on friendship networks), it is reasonable that they would be more likely to enhance life satisfaction.

This study suffers from methodological limitations. Some important limitations arise from our data. First, the outcome is based on a single question. Single-item variables are known to have lower reliability and validity than multi-item scales; however this is not necessarily the case in happiness [39]. In fact, previous attempts to validate multi-item happiness scales (e.g. PGC Morale Scale, SWLS, and PWI) have failed, and the correlates of the scale may be confounded with the scale itself [39].

A second limitation lies in the categorical nature of the outcome. Social desirability bias is common in happiness questions [126,199]. Thus, bounding responses to four categories (i.e. very unsatisfied; somewhat unsatisfied; somewhat satisfied; very satisfied) may result in a ceiling effect, whereby respondents prefer to be seen as doing better than just “somewhat satisfied”. This results also in small sample sizes for those who are less than “very satisfied”, which limits statistical inference. We partially overcame these limitations by dichotomizing the outcome (very satisfied vs. < very satisfied). In addition, ceiling effects are typical of happiness data across the world. For example, in the US 80%-85% of respondents report their lives as very satisfying or satisfying[39]. In other studies, such limitations could be partially overcome by asking respondents to rate their satisfaction numerically (i.e. from 1-10). Reporting life satisfaction as a number is more impersonal; this may ameliorate social desirability bias, whereby respondents may not be as reluctant to respond lower categories. It would also broaden the repertoire of possible responses, further alleviating the ceiling effect.

Changing interviewee moods may affect the reliability of happiness responses. However this tends to be less of a problem with life satisfaction questions, such as the ones used here [39]. Another limitation may lie in the exclusion of proxy interviews. In principle, we could have imputed the happiness of proxy respondents. First, it is risky to impute the outcome variable, even more so when imputations—as in the current case—would affect 25% of the sample (703 respondents out of 2,827). Second, proxy respondents are substantially different from the rest of the sample (older, suffer cognitive impairment, disability, and others). Imputation methods, which typically assume that data are missing at random, would provide unreliable approximations of their life satisfaction. Therefore, the exclusion of these observations is a superior alternative.

A further limitation lies in the "black box" phenomenon -- our analyses cannot quantify the causal chain between GC and life satisfaction. One example are mediator variables. Mediators stand in the causal chain and "affect the direction and/or strength of the relation between an independent… and a dependent…variable"[113]. Statistical tests are available to test for mediation; these assume that no endogeneity is present, and are difficult to model with non-continuous variables [202]. As some intermediate outcomes are likely to be endogenous (i.e. poverty variables), the corresponding mediation tests in Stata 11 (-sgmediation- and –binary-mediation- would yield inconsistent estimates. Therefore, efforts concentrated on removing endogeneity, rather than quantifying mediation. In addition, as the results for hypothesis 1 are inconclusive, mediation testing is not crucial. Finally, CRELES does not provide information on the intensity of program use. This precludes from establishing a dose-response relationship.

There are also study design limitations. The first concerns the effective removal of endogeneity. Though two important techniques were used to treat endogeneity (IV and RDD), there is never certainty of its complete removal. Second, the RDD analysis proceeded under acceptable, but not ideal conditions, which may have limited its capability to identify significant effects. In RDD, concerns may be raised about the fact that there are other discontinuities in the outcome at ages 62 and 76. Although these discontinuities are not as large at age 65 (the RDD cutoff), their presence violates one RDD assumptions [185]. On the other hand, the RDD model displays some strengths. There is an exogenous cutoff for program enrollment (i.e. 65 years old), and a reasonable amount of observations exists below and after the cutoff. Indeed, 390 observations are within 2 years of the 65 age cutoff. Further, as explained in table 2, the GC and non-GC groups are quite similar at baseline; our analyses compare groups that are relatively similar. Further, the results were consistent across bandwidths. Therefore, though still imperfect, we believe that our RDD models are reasonably valid.

The previous limitations do not warrant overlooking the study's strengths. First, this is one of the first program evaluations to use happiness as an outcome. Second, we were able to find a strong and valid instrument for our IV analyses. Third, the wide variety of analyses (DiD, LDVM, change-scores) fortify the study's findings. Key amongst these is DiD, which cancels out time-invariant unobservable factors. Fourth, the use of a large survey permits overcoming previous sample size concerns, which had pervaded happiness literature of elderly populations.

Finally, this experience points to crucial areas for future evaluations. First, researchers should use whenever possible a continuous measure of happiness. This should alleviate sample size concerns, and partially mitigate the ceiling effects in life satisfaction. Second, respondents should be asked to rate their happiness both globally and across domains (i.e. how satisfied they are with their family life, their job, and so on). This is because there are indications that overall happiness does not necessarily equal the sum of its components [1]. Third, researchers should use validated happiness scales. Some scales are available, such as the Subjective Happiness Scale[203], the Satisfaction with Life Scale[204], and the Positive and Negative Affect Scale[205]. Fourth, it is essential to obtain detailed program exposure information, including frequency and intensity of use. Such information would permit determining dose-response relationships, a key aspect of causality. Fifth, mixed-method designs can strengthen significantly the interpretability of the findings. By including qualitative research, such designs can shed light on aspects that are sometimes difficult to quantify, such as the extent of program implementation, and the quality of the intervention.

Conclusion

There are sound reasons for using happiness as an evaluation criterion. First, the happiness phenomenon has occupied classical and liberal thinking, from Aristotle to Jeremy Bentham [12]. Indeed, it has been claimed that happiness is "the one thing that everybody wants from life"[2]. Second, though happiness may not be the ultimate life goal for all (see [122]); even skeptics assume it as a key ingredient of alternative constructs, such as the good life[206] in [207], p.3. Third, governments in different continents (France, United Kingdom, Bhutan) are increasingly using happiness as a policy outcome. Fourth, including happiness in program evaluations can help elucidate whether government policy can influence people's happiness. By doing so, policy-makers can understand which policy instruments impact what truly matters to people, and channel investments more efficiently. By doing so, happiness could realign policy-making in profound ways. This would entail shifting the emphasis away from GDP growth, which does not necessarily improve life satisfaction, and levies heavy environmental costs. Such ideas were summarized eloquently by Robert Kennedy[208].

"Our gross national product ... counts air pollution and cigarette advertising, and ambulances to clear our highways of carnage. It counts special locks for our doors and the jails for those who break them. It counts the destruction of our redwoods and the loss of our natural wonder in chaotic sprawl.

Yet the gross national product does not allow for the health of our children, the quality of their education, or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages; the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage; neither our wisdom nor our learning; neither our compassion nor our devotion to our country; it measures everything, in short, except that which makes life worthwhile".

In addition, happiness could become the common metric to evaluate disparate government programs. It is currently difficult to compare competing public programs, as they use sector-specific intermediate outcomes. How to decide if a five point reduction in myocardial infarctions is preferable to attaining a

5% improvement in math scores? The intermediate outcomes of such interventions are not comparable; it is therefore impossible to decide which one is preferable based solely on intermediate outcomes. By using a common underlying construct, these programs can be compared head-to-head, and resources can be allocated to where they matter the most.

Promoting happiness (through program evaluations and other measures) could generate positive downstream externalities. For example, increased happiness can harness social capital [209], which can revert favorably on happiness itself [177]. Monitoring happiness at the societal level can also reveal important information to the policy maker, such as social dysfunction and depression [209], and help identify those groups most likely to benefit from public policy. Using 17 years of data, Frank Fujita and Ed Diener found that up to a quarter of the population can change their happiness levels over time[210]. Identifying these individuals may prove crucial to sound happiness policy.

There are signs that the "happiness culture" is taking root amongst high-level academics and policy-makers. Last year the Earth Institute published the first World Happiness Report[27]; in a similar note, the Legatum Institute has developed a prosperity index to compare countries. This index is based on national income and life satisfaction [211]. Through such initiatives, research could contribute to happier societies, and higher value-added government investments. The insights from this experience, and many others to follow, may eventually contribute to sound happiness policy.

Tables

Table 1: Population Characteristics, by CRELES Round

	Period 1 (n=2,011)	Period 2 (n=1,684)	Change	p-value
Very satisfied (%)	74.9	78.2	3.3	*0.041*
Life satisfaction score	3.7	3.7	0.0	*0.050*
Age (years)	69.3	70.4	1.1	*0.000*
Sex (% female)	51.8	52.5	0.7	*0.934*
Education (nr. years)	5.5	N/A	N/A	*N/A*
In a relationship (% married or cohabiting)	63.4	62.1	-1.3	*0.472*
Religious frequency (% go to church >=1 time/weekly)	53.3	55.4	2.1	*0.271*
Intense contact with children (% sees children >=1 time/weekly)	78.7	80.4	1.7	*0.284*
Exercises (% exercises >=3 times/week)	34.8	33.4	-1.4	*0.449*
Income (log thousands colones, nominal)	4.4	4.6	0.2	*0.002*
Poverty (% below poverty line)	13.8	12.6	-1.3	*0.294*
Assets	13.9	14.3	0.4	*0.112*
Locus of control score	3.1	N/A	N/A	*N/A*
Depression (% depressed)	6.0	5.6	-0.4	*0.675*
Household composition (nr of people in household)	3.4	3.2	-0.2	*0.010*
Urban residence (% urban)	63.6	62.9	-0.7	*0.694*

** Education and Locus of Control were only asked in period 1*
** Analyses take sampling weights into account*

Table 2: Differences in Life Satisfaction Predictors, Golden Citizen Members vs. Non-Members

CRELES, round 1				
	GC (n=1,427)	**non-GC (n=683)**	**Difference**	**p-value**
Very satisfied (%)	74.5	75.1	-0.6	*0.311*
Age (years)	73.2	67.8	5.4	*0.000*
Sex (% female)	49.6	52.5	-2.9	*0.280*
Civil status				
In a relationship (%)	56.6	65.6	-8.9	*0.000*
Widow (%)	31.8	20.2	11.5	*0.000*
Single (%)	6.6	7.3	-0.7	*0.613*
Divorced (%)	5.2	5.0	0.2	*0.884 1*
Years of education	5.6	5.5	0.1	*0.784*
Pensioned (%)	68.3	44.4	23.8	*0.000*
Health status				
>=3 chronic diseases (%)	30.5	27.3	3.2	*0.179*
Disabled functionally (%)	13.7	9.0	4.7	*0.003*
Urban residence (%)	67.9	62.1	5.8	*0.020*
Depressed (+10 YGDS scale; %)	4.7	6.4	-1.7	*0.155*
Goes to church > 1 week (%)	53.0	53.3	-0.3	*0.901*
Sees children regularly (%)	80.3	78.7	1.6	*0.463*
Number of household members (count)	3.2	3.4	-0.2	*0.015*
Personality				
Locus of control (score)	3.2	3.1	0.1	*0.423*

** Analyses take sampling weights into account*

Table 3: Differences in Intermediate outcomes, Golden Citizen Members vs. Non-Members, Round 1

	GC (n=1,427)	Non-GC (n=683)	Difference	p-value
Waits < 30 min at doctor's office (%)	48.3	43.4	4.9	*0.065*
Visited by EBAIS team in last 6 months (%)	44.2	40.4	3.8	*0.149*
Average income (thousand colones)	138.2	130.2	8.1	*0.659*
Below poverty line (%)	12.2	14.4	-2.2	*0.203*
Below extreme poverty line (%)	3.7	6.7	-3.0	*0.007*
Assets (points)	14.2	13.8	0.4	*0.248*
Exercising				
Exercises >=3 times a week (%)	28.6	37.0	-8.3	*0.001*
Nr. weekly hours of exercise (count)	2.5	3.7	-1.2	*0.013*
Social networks				
Nr of friends (count)	3.7	3.7	-0.0	*0.796*
Has at least 1 friend (%)	97.6	97.1	0.5	*0.571*
Talks to friends at least once per week (%)	95.0	84.8	10.2	*0.940*

** Analyses take sampling weights into account*

Table 4: Associations between Life Satisfaction and Intermediate Outcomes, Round 1

	Very satisfied	< Very satisfied	Difference	p-value
Healthcare access				
Waits < 30 min at doctor's office	46.7	38.3	8.4	*0.005*
Was visited by EBAIS team in last 6 months (%)	42.4	38.3	4.1	*0.163*
Income				
Average income (thousand colones)	150.0	78.5	71.5	*0.000*
Below poverty line (%)	11.4	21.0	-9.5	*0.002*
Below extreme poverty line (%)	4.3	10.8	-6.5	*0.000*
Assets (points)	14.4	12.2	2.2	*0.000*
Exercising				
Exercises >=3 times a week (%)	36.9	28.8	8.0	*0.006*
Nr. weekly hours of exercise (count)	3.5	3.2	0.3	*0.665*
Social networks				
Nr of friends (count)	3.8	3.5	0.3	*0.000*
Has at least 1 friend (%)	97.7	94.9	2.8	*0.007*
Talks to friends at least once per week (%)	85.8	80.7	5.1	*0.037*

** Analyses take sampling weights into account*
** EBAIS are the primary care centers of the public Costa Rican health service*
** Social relationship variables only available at wave 2*
** Sample restricted to respondents with information in both waves*

Table 5: Propensity Scores - Mean Overall Standardized Bias

	Mean standardized bias	
Propensity scores	**Raw**	**Matched**
Simple model, round 1	7.1	0.2
Simple model, round 2	17.5	4.8

** Simple model propensity scores*

Table 6: Changes in Golden Citizen Membership, Rounds 1 and 2

Membership status		
Period 1	**Period 2**	**%**
Not enrolled	Not enrolled	51.0%
Enrolled	Not enrolled	6.0%
Not enrolled	Enrolled	26.0%
Enrolled	Enrolled	17.1%

** Analyses take sampling weights into account*
** A chi-2 test of proportions was significant (p<0.001)*

Table 7: Multivariate Estimations of the Relationship between Golden Citizen Status and Life Satisfaction

	X=(Golden Citizen)	
	Marginal coefficient	**Standard Error**
Main analyses		
Instrumental variables (IV)		
Probit (Very satisfied)	-4.2%	*0.008***
Regression Discontinuity Design (RDD)		
Primary analysis		
Y= Probit(Very satisfied | cond prob GC=0.1), bandwith 1.7	-7.7%	*0.228*
Y= Probit(Very satisfied | cond prob GC=0.1), bandwith 0.85	-7.9%	*0.229*
Y= Probit(Very satisfied | cond prob GC=0.1), bandwith 3.4	-2.8%	*0.304*
Alternative bandwidths in RDD		
Y= Probit(Very satisfied | cond prob GC=0.3)	-20.0%	*0.594*
Y= Probit(Very satisfied | cond prob GC=0.5)	-77.4%	*2.925*
Y= Probit(Very satisfied | cond prob GC=0.9)	12.0%	*0.350*
Sensitivity analyses		
Difference in differences (DiD)		
Unadjusted		
Very satisfied	-2.5%	*0.033*
Propensity-score adjusted (short form)		
Very satisfied	0.4%	*0.038*
Lagged Dependent Variable Model (LDVM; Y2 on X2 and Y1)		
Unadjusted		
Very satisfied	-5.4%	*.023***
Propensity-score adjusted (short form)		
Very satisfied	1.1%	*0.024*
Change-scores (ΔY on X1; 4 categories)		

Unadjusted		
Persistent dissatisfaction	6.1%	*0.015***
Satisfaction decrease	0.4%	*0.015*
Satisfaction increase	5.2%	*0.015***
Persistent satisfaction	-11.6%	*0.023****
Propensity-score adjusted (short form)		
Persistent dissatisfaction	5.2%	*0.016***
Satisfaction decrease	-0.3%	*0.016*
Satisfaction increase	6.6%	*0.017****
Persistent satisfaction	-11.5%	*0.024****

** In change-scores models, the outcome is modeled as follows: persistent dissatisfaction (< very satisfied at both time points); satisfaction decrease (very satisfied - < very satisfied); satisfaction increase (<very satisfied - very satisfied); persistent satisfaction (very satisfied at both timepoints).*
** Propensity scores: 1:1 PSM for LDVM and change-scores; weights by odds for DID analysis*
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Table 8: Multivariate Estimations of the Relationship between Golden Citizen Membership and Intermediate Outcomes

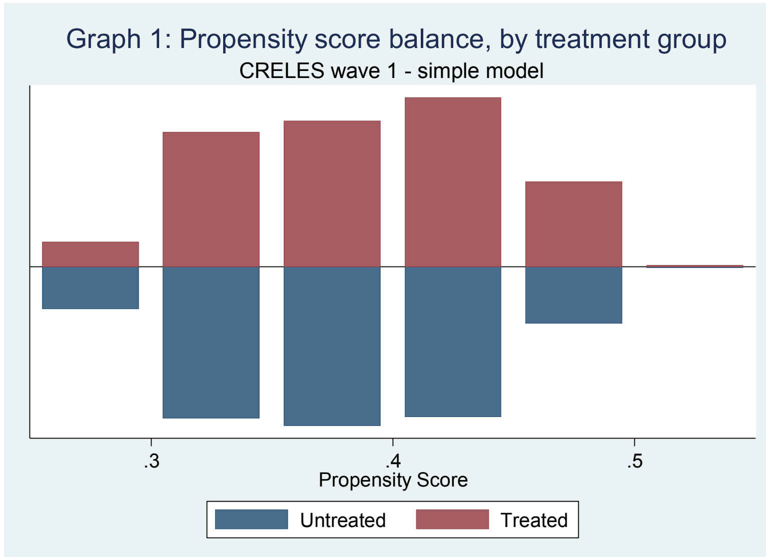
	X=Golden Citizen	
	Marginal coefficient	**Standard Error**
Instrumental variables (IV)		
Healthcare access		
< 30 min wait at doctor office	2.1%	0.005**
>1 EBAIS visit in last 6 months	1.9%	0.026
Log income	-3.5%	0.195
Poverty		
Poor	-8.3%	.032**
Extremely poor	-5.6%	.027**
Exercise		
Exercises >=3 times a week (%)	1.2%	0.008
Social networks		
Nr of friends (count)	-0.15	0.185
Has at least 1 friend (%)	-2.7%	0.042
Talks to friends at least once per week (%)	3.9%	0.008**
Regression Discontinuity (RDD)		
Y=Probit(Channel)		
Healthcare access		
< 30 min wait at doctor office	39.9%	0.350
>1 EBAIS visit in last 6 months	-22.6%	0.331
Poverty		
Poor	-13.7%	0.284
Extremely poor	-5.8%	0.159
Exercise		
Exercises >=3 times a week (%)	-17.0%	0.368
Social networks		
Nr of friends (count)	0.6	0.012
Has at least 1 friend (%)	-2.9%	0.064
Talks to friends at least once per week (%)	1.3%	0.029
Difference in differences		
< 30 min wait at doctor office	0.7%	0.037
>1 EBAIS visit in last 6 months	-2.5%	0.036
Log Income	0.3	0.087**
Poor	-6.1%	0.030**

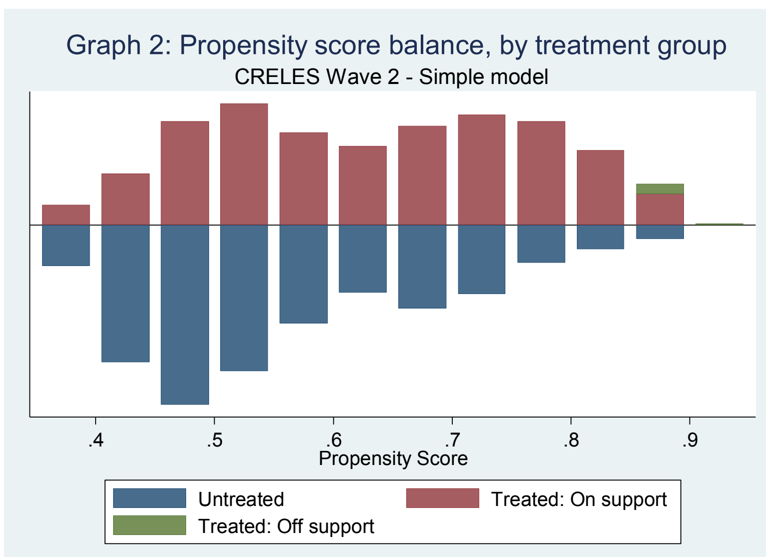
Extremely poor	-5.3%	0.022**
Lagged Dependent Variable Model (LDVM; Y2 on X2 and Y1)		
< 30 min wait at doctor office	-2.4%	0.028
>1 EBAIS visit in last 6 months	1.8%	0.028
Poor	-5.8%	.017 **
Extremely poor	-2.2%	0.001**
Exercise	6.0%	0.026 **
Social networks		
Nr of friends (count)	1.4%	0.008 *
Has at least 1 friend (%)	2.6%	0.021
Talks to friends at least once per week (%)	3.6%	0.042
Change-scores (Y2-Y1 on X1)		
< 30 min wait at doctor office		
No-No	-4.9%	0.024**
Yes-No	-0.2%	0.001
No-Yes	1.2%	0.006**
Yes-Yes	3.9%	0.019**
>1 EBAIS visit in last 6 months		
No-No	1.7%	0.025
Yes-No	0.0%	0.000
No-Yes	-0.3%	0.004
Yes-Yes	-1.5%	0.021
Poor		
No-No	7.0%	0.022 **
Yes-No	-1.8%	0.006**
No-Yes	-1.9%	0.006**
Yes-Yes	-3.4%	0.011**
Extremely poor		
No-No	7.2%	0.018***
Yes-No	-3.3%	0.009***
No-Yes	-2.3%	0.007***
Yes-Yes	-1.6%	0.005**
Exercise		
No-No	1.2%	0.025
Yes-No	-0.2%	0.003
No-Yes	-0.3%	0.006
Yes-Yes	-0.7%	0.015
Social networks		
Nr of friends (count)	0.0	0.042

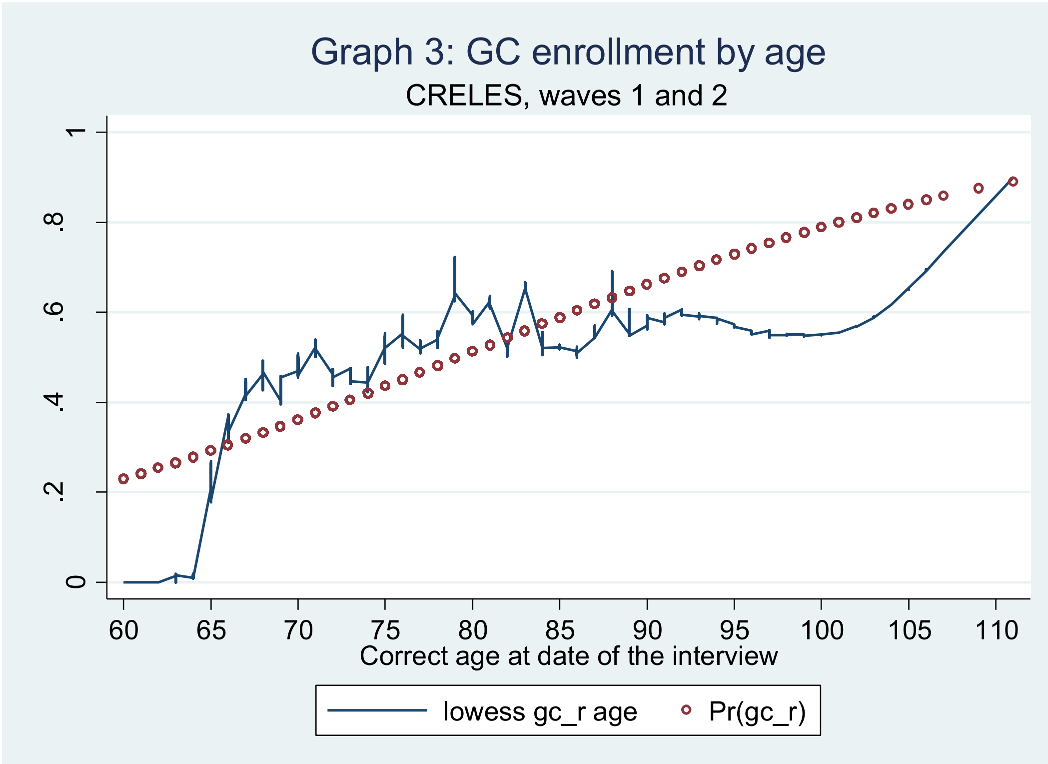
Has at least 1 friend (%)	1.5%	0.009 *
Talks to friends at least once per week (%)	2.80%	0.021

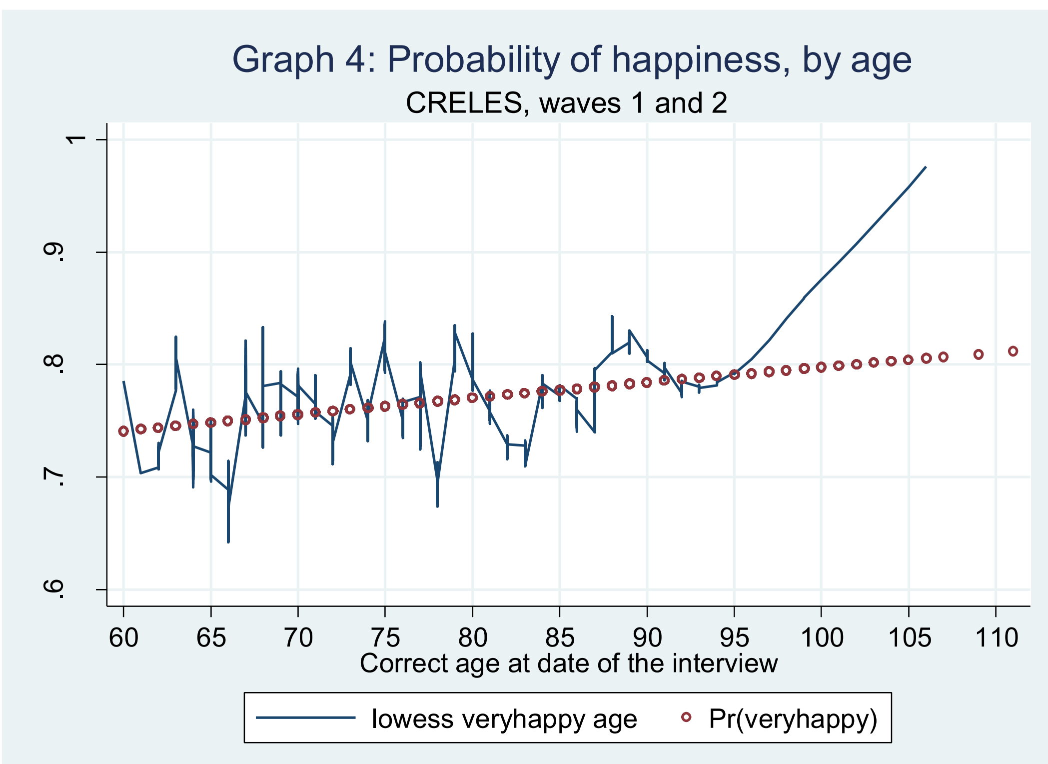
** Standard error legends: * if < 0.1; ** if <0.05; *** if < 0.001*

Graphs


Graph 1: Propensity score balance, by treatment group
CRELES wave 1 - simple model
.3
.4
.5
Propensity Score
Untreated
Treated


Graph 2: Propensity score balance, by treatment group
CRELES Wave 2 - Simple model
.4
.5
.6
.7
.8
.9
Propensity Score
Untreated
Treated: On support
Treated: Off support


Graph 3: GC enrollment by age
CRELES, waves 1 and 2
0
.2
.4
.6
.8
1
60
65
70
75
80
85
90
95
100
105
110
Correct age at date of the interview
lowess gc_r age
Pr(gc_r)


Graph 4: Probability of happiness, by age
CRELES, waves 1 and 2
.6
.7
.8
.9
1
60
65
70
75
80
85
90
95
100
105
110
Correct age at date of the interview
lowess veryhappy age
Pr(veryhappy)

5. CONCLUSION

This thesis explores whether health policy can influence happiness. An understanding of this relationship sheds light on whether governments can promote happiness.

These questions are relevant to policy makers and researchers. As discussed in the introduction, both national governments (France, United Kingdom, and others) and international organizations (European Union, OECD) have adopted happiness as an explicit policy goal, and are collecting data to measure its achievement. This signals a paradigm shift, which recognizes that government policy must improve the whole human experience. In the same tune, researchers are shedding light on the components of the happiness construct, its measurements, and etiology.

All these developments are laying the foundations for a new happiness policy[208]. However, gaps have persisted. It has been unclear which health dimensions are most important to happiness; the effect of childhood health on adult happiness has been seldom evaluated; perhaps more importantly, happiness has been seldom used as an evaluation yardstick. This doctoral thesis aimed to close these gaps by considering the following questions:

- Is it subjective or objective health that affects happiness?
- Is childhood health related with happiness in late life?
- Has a government program (Golden Citizen) affected happiness and its intermediate outcomes (healthcare access, poverty status, exercising, and social relationships)?

The conclusion chapter summarizes the empirical findings, discusses the research and policy implications, and summarizes the lessons learned in this effort.

Discussion of the empirical evidence

The main findings of this thesis are:

- Subjective health in more important than objective health in determining happiness;
- Childhood adversity may influence happiness throughout the life course;
- Governmental interventions that improve objective conditions may not impact happiness.

The common lesson is that impacting objective conditions does not necessarily affect happiness. Therefore, it is possible that many public policies and programs, which focus on improving objective circumstances, will not necessarily improve happiness.

The first article shows that objective measures of health may not influence happiness. Policy-makers need therefore a better understanding of which specific components of subjective health are related to happiness. Depression treatment and pain management are potential venues. Anxiety and other dimensions of mental health not addressed in this study could be also acted upon in order to affect happiness. The second essay underlines that childhood health appears to influence happiness --- but mainly in its subjective version (i.e. having experienced bad health during the first 15 years of life). In fact, in line with previous research[58], most objective measures of childhood health (specific conditions) showed no significant relationship with happiness. In the same vein, it becomes necessary to understand which subjective components of childhood health are related to happiness, and how to influence them.

This study was able to assess only an overall self-reported measure; it would be important to assess other measures (pain, depression, etc. during childhood) and their determinants. Through these, policy-makers may be able to affect the wellbeing of individuals throughout the lifecourse. The last article shows that well-intentioned programs for the elderly, which improve objective conditions such as poverty, may not necessarily succeed in improving happiness. Therefore, impacting intermediate outcomes does not equate impacting wellbeing. If programs are assumed to affect happiness, then they must be also evaluated in terms of happiness. That way, policy-makers could design programs that impact their intended bottom line.

Taken together, these findings suggest a potential (or partial) disconnect between objective life circumstances and happiness. This may occur because individuals adapt; because subjective assessments predominate over objective life conditions; or because the interventions that ameliorate objective conditions also carry negative consequences that outweigh their benefits. On the adaptation front, it is important to acknowledge the elderly nature of our sample. Many interviewees may have been suffering adverse health for a while; since times of diagnosis are not available, they may have adapted to their conditions by the time of the survey interview. This is even more likely with conditions that happened far back in childhood. There is also no information on when individuals receive the Golden Citizen interventions; therefore, they may have also adapted by the time of their interview. The overriding importance of subjective predictors calls for a finer understanding of the happiness determinants. For example, the main adult health predictors of happiness in this study were all subjective (self-reported overall health; depression; pain). If health policy is to truly affect happiness, these “soft” (subjective) targets are more important than the “hard” (objective: blood pressure, etc.) health targets. A similar discourse applies to childhood health determinants of happiness, which in this thesis also turn out to be subjective. Finally, it is possible for the negative costs of interventions to outweigh their benefits. For example, the improved health outcomes of disease control programs may positively impact happiness; yet, their component activities could entail personal costs (constant visit to the hospital, financial worries, etc.) that are detrimental to these gains. Similarly, Golden Citizen program may levy some personal costs (enrollment complicacies, bureaucratic complexities, and so on) with adverse effects on life satisfaction.

This is not to say that such interventions do not have value –indeed, public interventions often do impact their intermediate outcomes. But in order to impact happiness, these programs may need to redirect their goals, or include new components explicitly targeted at improving happiness. These are likely to be of “soft” (subjective) nature. For example, since—as found in this study—mental health conditions have a principal effect on happiness, they should get more attention from health policy. For childhood health policy to truly affect happiness, policy-makers need to act upon the dimensions of childhood subjective health that affect happiness. In the Golden Citizen program, it is possible that improving poverty is a necessary but insufficient condition to improve the happiness. If improving happiness is indeed a program goal, policy-makers could consider fortifying it with additional components, such as those of non-pecuniary nature, which seem to have long lasting effects on happiness[34,153]. For example, it could be thought that the social component of Golden Citizen could bring substantial benefits in terms of happiness.

Strengths and limitations

The previous statements are qualified some study limitations. These are mentioned in the individual chapters; we discuss here only those that are common to all three articles. The first limitation lies in the life satisfaction question in our survey. Responses were coded as categorical (four levels); as discussed previously, constraining responses to four levels, and the wording of the questions may have compelled participants to overstate their happiness. A continuous, numeric form would have given more response options, and de-personalize the answers , thereby reducing response biases (i.e. responding to feeling a happiness of 7 in a sale of 1-10 can be less negatively viewed than "somewhat happy"). In such case, the outcome would exhibit a milder ceiling effect and higher variability, both of which improve statistical inference. Another limitation lies in the fact that our survey is limited to the elderly. Consequently, findings can be only generalized to this segment of the population. However, as mentioned throughout the article, the senior population of Costa Rica (and indeed all of Latin America) is projected to grow rapidly. Research on the wellbeing of seniors is important at this inflection stage. Another criticism could be that CRELES contains limited information on previous life stages. It is possible that life satisfaction in seniors is product of continued exposure to health and other factors; yet, many of these (particularly those fixed over time) have been adjusted for through econometric techniques (IV, RDD, DiD, LDVM and change-scores). Finally, the survey did not contain many variables that could have been used as instruments. We found an instrument only for our third article (evaluation of the Golden Citizen program); this methodological weakness was partially overcome through a wide variety of sensitivity analyses.

Implications for research

Happiness research is expanding rapidly. However, the study of happiness is still new in public health. This stands spite the fact that happiness is integral to health: health is a state of complete wellbeing[35], of which happiness is a necessary component[60]. The literature review revealed a few studies that explored the effect of health interventions on happiness [212]; however these have been seldom carried out by public health researchers. Given the inter-disciplinarity of public health, and that happiness is a complex phenomenon that requires analysis from competing perspectives, public health researchers are uniquely suited for happiness research.

A few areas stand out for future research. First, a deeper understanding of the relationships between mental health and happiness is to be gained. Other than depression, this study was not able to assess the mental health conditions that affect happiness. We were also not able to assess whether the length of disease or disease intensity are associated with happiness; this may be particularly important for chronic diseases. Second, it is important to gain a better understanding of how childhood subjective health status (mental health, pain and others) could affect happiness. Third, lifecourse studies are critical to establish the influence of childhood health on happiness. Such studies could help understand whether the effect of childhood health on happiness varies across time[145], and whether childhood health (in combination with genetics) could contribute to establishing a happiness setpoint early in life. Third, program evaluations should use happiness as a standard outcome. These should use econometric techniques (such as IV, RDD and DiD) to remove endogeneity. Fourth, the role of adaptation warrants further research. Adaptation implies that individuals revert to their fixed happiness setpoint after major life changes. There is evidence that the happiness of individuals adapts to some circumstances, but not to others[153]. A few studies have studied adaptation to certain health conditions[49,117]; however it is still unclear which health dimensions and interventions have long lasting effects on happiness. Such research could shed light on important policy questions such as whether people adapt to some health conditions and not others; the length and quality of adaptation; whether adaptation is related to disease severity, and which interventions may favor—or preclude—adaptation. Understanding these relationships is critical to attaining health interventions with the highest potential on wellbeing. Fifth, the interactions of health with other life dimensions (poverty, social relationships, and others) are not known. While it is possible that a specific intervention (i.e. healthcare access) per se may not impact happiness, its interaction with another dimension (i.e. poverty) may result in improved wellbeing. These interactions have been seldom discussed, and should become part of future endeavors.

Finally, sound happiness research hinges on the development of a valid, reliable happiness scale. Currently no such scale exists; this may be therefore be the most important milestone in happiness research in the short term. A prerequisite for this is the attainment of a consensus on the happiness concept. Such situation is not uncommon; depression was also initially a blurry construct, which became clearer over time[213]. A standard measure would integrate the competing happiness concepts (positive affect, life satisfaction, and so on) into a single measure. Ideally, such metric would have international validity. This would allow inter-country comparisons, which is currently a challenge in happiness research[126]. The efforts to this end by the National Science Foundation

[10] will establish happiness research as an important area of inter-disciplinary activity.

Implications for policy

The current effort suggests also some policy implications. First, as mentioned at the beginning of this chapter, it is possible for programs to affect intermediate outcomes but not on happiness. If happiness is indeed a program goal, policies must be also evaluated in terms of happiness. The development of a happiness scale would be fundamental to this end. This way, policy-makers could gain a better understanding on which interventions contribute to happiness. Most importantly, a standard happiness metric would allow the head-to-head comparison of from different sectors with a valid, single measure. This would be a key benefit to policy-makers, who are currently unable to compare interventions from different sectors. The existence of a validated happiness metric would boost its validity as an overt policy goal, which could facilitate the integration of various public sectors (health, education, etc.). Such integration would help overcoming inter-departmental silos, which are currently pervasive in public administrations. Measuring and using happiness in policy can also suggest implementation difficulties or areas for improvement. For example, if a program has a positive impact on an intermediate variable but not on happiness, then unintended effects (i.e. negative externalities) may be present. In line with earlier examples, a program may enhance friendship networks, but to the expense of family relationships. Such a situation could raise a red flag for the policy-maker.

Although this doctoral thesis has emphasized the importance of happiness as a policy goal, it is worth noting the risks of an overemphasis on happiness. Self-development, justice and freedom of choice are also fundamental to the human experience[214]. It has been found that individuals are willing to sacrifice their happiness at the behest of high ideals, such as altruism and fairness [36]. Hence, happiness may be a necessary, but insufficient indicator of progress. This is the underlying philosophy of the newly developed prosperity indicators of the OECD´s Better Life Index[215], Thailand´s Green and Happiness Index, and the Gross National Happiness Index of Bhutan (both at[216]). These composite indicators emphasize other life dimensions as well, such as education, income, safety, civil engagement, ecological quality, justice, and good governance. The use of such indices can result in integrated public policies that impact human experience as whole.

Conclusion

Public programs assume that they will increase happiness by improving objective life dimensions. However, this thesis has shown that an exclusive focus on these indicators may not necessarily improve happiness. For policies and programs to be truly effective to this end, they must gain a better understanding of the happiness determinants; public programs must be evaluated in terms of happiness (in addition to intermediate outcomes), and the prosperity of countries should be based on comprehensive indicators that include happiness, wealth, and other important life dimensions. Through these actions, governments can realign their policies to impact their ultimate goal –the pursuit of human happiness.

REFERENCES

1. Diener E, Napa-Scollon CK, Oishi S, Dzokoto V, Suh EM. Positivity and the construction of life satisfaction judgments: Global happiness is not the sum of its parts. *Journal of happiness studies*. 2000;1(2):159-176.

2. Fordyce MW. A program to increase happiness: Further studies. *Journal of Counseling Psychology*. 1983;30(4):483.

3. Ebrecht M, Hextall J, Kirtley L, Taylor A, Dyson M, Weinman J. Perceived stress and cortisol levels predict speed of wound healing in healthy male adults. *Psychoneuroendocrinology*. 2004;29(6):798-809.

4. Cohen S, Alper CM, Doyle WJ, Treanor JJ, Turner RB. Positive emotional style predicts resistance to illness after experimental exposure to rhinovirus or influenza A virus. *Psychosom Med*. 2006;68(6):809-815.

5. Graham C, Eggers A, Sukhtankar S. Does happiness pay?:: An exploration based on panel data from russia. *Journal of Economic Behavior & Organization*. 2004;55(3):319-342.

6. Seligman MEP. Positive psychology, positive prevention, and positive therapy. *Handbook of positive psychology*. 2002;2:3-12.

7. Dutt AK, Radcliff B. *Happiness, economics and politics: Towards a multi-disciplinary approach.* Edward Elgar Publishing; 2009.

8. Self A, Thomas J, Randall C. Measuring national well-being: Life in the UK, 2012. . 2012.

9. Stiglitz JE, Sen A, Fitoussi JP. Report by the commission on the measurement of economic performance and social progress": Commission on the measurement of economic performance and social progress. . 2009.

10. de Vos M. Saving happiness from politics. *National Affairs*. 2012.

11. Helliwell J, Layard R, Sachs J(. World happiness report. . 2012.

12. Graham C. The economics of happiness. *World Economics*. 2005;6(3):41-55.

13. Russell B. *The conquest of happiness.* Routledge; 2006.

14. Seligman MEP, Csikszentmihalyi M. Positive psychology: An introduction. *American Psychologist; American Psychologist*. 2000;55(1):5.

15. Inglehart R, Foa R, Peterson C, Welzel C. Development, freedom, and rising happiness: A global perspective (1981–2007). *Perspectives on psychological science*. 2008;3(4):264-285.

16. Howell RT, Howell CJ. The relation of economic status to subjective well-being in developing countries: A meta-analysis. *Psychol Bull*. 2008;134(4):536.

17. Johnson W, Krueger RF. How money buys happiness: Genetic and environmental processes linking finances and life satisfaction. *J Pers Soc Psychol*. 2006;90(4):680.

18. Gardner J, Oswald A. Does money buy happiness? A longitudinal study using data on windfalls. *Warwick University mimeograph*. 2001.

19. Gardner J, Oswald AJ. Money and mental wellbeing: A longitudinal study of medium-sized lottery wins. *J Health Econ*. 2007;26(1):49-60.

20. Di Tella R, New JHD, MacCulloch R. Happiness adaptation to income and to status in an individual panel. *Journal of Economic Behavior & Organization*. 2010.

21. Stutzer A. The role of income aspirations in individual happiness. *Journal of Economic Behavior & Organization*. 2004;54(1):89-109.

22. Brickman P, Campbell DT. Hedonic relativism and planning the good society. *Adaptation-level theory*. 1971:287–305.

23. Brickman P, Coates D, Janoff-Bulman R. Lottery winners and accident victims: Is happiness relative?. *J Pers Soc Psychol*. 1978;36(8):917.

24. Easterlin RA, McVey LA, Switek M, Sawangfa O, Zweig JS. The happiness–income paradox revisited. *Proceedings of the National Academy of Sciences*. 2010;107(52):22463.

25. Easterlin RA. Does money buy happiness? *Public Interest*. 1973;30(3):3-10.

26. Diener E, Lucas RE, Oishi S, Suh EM. Looking up and looking down: Weighting good and bad information in life satisfaction judgments. *Person Soc Psychol Bull*. 2002;28(4):437-445.

27. Helliwell J, Layard R, Sachs J. World happiness report. . 2012.

28. Kahneman D, Deaton A. High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*. 2010;107(38):16489-16493.

29. Prentice AM. The emerging epidemic of obesity in developing countries. *Int J Epidemiol*. 2006;35(1):93-99.

30. Beekman AT, de Beurs E, van Balkom AJ, Deeg DJ, van Dyck R, van Tilburg W. Anxiety and depression in later life: Co-occurrence and communality of risk factors. *Am J Psychiatry*. 2000;157(1):89-95.

31. Global Footprint Network. The national footprint accounts, 2011. . 2012.

32. Alliance Development Works. World risk report 2012. . 2013.

33. The World Bank. Environment: At a glance. http://web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:20036126~menuPK:34480~pagePK:34370~theSitePK:4607,00.html. Accessed 05/05, 2013.

34. Layard PRG, Layard R. *Happiness: Lessons from a new science.* Penguin Group USA; 2006.

35. World Health Organization. *Constitution of the world health organization: Signed at the international health conference, new york, 22 july 1946.* World Health Organization, Interim Commission; 1947.

36. Powdthavee N. Economics of happiness: A review of literature and applications. *Chulalongkorn Journal of Economics*. 2007;19(1):51-73.

37. Lykken D, Tellegen A. Happiness is a stochastic phenomenon. *Psychological Science*. 1996;7(3):186-189.

38. Dolan P, Peasgood T, White M. Do we really know what makes us happy A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*. 2008;29(1):94-122.

39. George LK. Still happy after all these years: Research frontiers on subjective well-being in later life. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2010;65(3):331.

40. Di Tella R, MacCulloch R. Some uses of happiness data in economics. *The Journal of Economic Perspectives*. 2006;20(1):25-46.

41. Lelkes O. Knowing what is good for you:: Empirical analysis of personal preferences and the. *J Socio-econ*. 2006;35(2):285-307.

42. Shields MA, Price SW. Exploring the economic and social determinants of psychological well-being and perceived social support in england. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2005;168(3):513-537.

43. Gerstenbluth M, Rossi M, Triunfo P. Felicidad y salud: Una aproximación al bienestar en el río de la plata. *Estudios de economía*. 2008;35(1):65-78.

44. Graham C, Higuera L, Lora E. Which health conditions cause the most unhappiness? *Health Econ*. 2011.

45. CId A, Ferrés D, Rossi M. Subjective well-being in the southen cone: Health, income and family. . 2008.

46. Lucas RE, Clark AE, Georgellis Y, Diener E. Reexamining adaptation and the set point model of happiness: Reactions to changes in marital status. *J Pers Soc Psychol*. 2003;84(3):527.

47. Lucas RE. Time does not heal all wounds. *Psychological Science*. 2005;16(12):945.

48. Calvo E, Haverstick K, Sass SA. Gradual retirement, sense of control, and retirees' happiness. *Res Aging*. 2009;31(1):112.

49. Oswald AJ, Powdthavee N. Does happiness adapt? A longitudinal study of disability with implications for economists and judges. *Journal of Public Economics*. 2008;92(5-6):1061-1077.

50. Winkelmann L, Winkelmann R. Why are the unemployed so unhappy? evidence from panel data. *Economica*. 1998;65(257):1-15.

51. Santana MS, Chaves Maia EM. Atividade física e bem-estar na velhice. *Revista de Salud Pública*. 2009;11(2):225-236.

52. Guedea MTD, Albuquerque FJB, Tróccoli BT, Noriega J, Seabra M, Guedea RLD. Relação do bem-estar subjetivo, estratégias de enfrentamento e apoio social em idosos. *Psicologia: Reflexão e Crítica*. 2006;19(2):301-308.

53. Senoo K, Takagi O. Helping behavior and psychosocial well-being in elderly people. *Shinrigaku Kenkyu*. 2004;75(5):428-434.

54. Okamoto K. Feeling of well-being and family contacts in community elderly residents. *Nihon Ronen Igakkai Zasshi*. 2000;37(2):149-154.

55. Nobe M.
The sense of well-being of elderly women in a medium-size japanese city. *Sociological Theory and Methods*. 1999:121-123.

56. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. *J Health Econ*. 2005;24(2):365-389.

57. Power C, Manor O, Fox J. *Health and class: The early years.* Chapman & Hall; 1991.

58. Frijters P, Johnston D, Shields M. Destined for (un) happiness: Does childhood predict adult life satisfaction? *IZA Discussion Paper No.5819*. 2011.

59. Population Division, Department of Economic and Social Affairs. World population prospects: The 2006 revision. . 2006.

60. Diener E. Subjective well-being: The science of happiness and a proposal for a national index. *Am Psychol*. 2000;55(1):34.

61. Bates W. Gross national happiness. *Asian-Pacific Economic Literature*. 2009;23(2):1-16.

62. Blanchflower DG, Oswald AJ. Is well-being U-shaped over the life cycle? *Soc Sci Med*. 2008;66(8):1733-1749.

63. Blanchflower DG, Oswald A. Well-being over time in britain and the USA. . 2000.

64. Clark A. Born to be mild? cohort effects don't (fully) explain why well-being is U-shaped in age. *IZA Discussion Paper No.3170*. 2007.

65. Graham C, Felton A. Inequality and happiness: Insights from latin america. *Journal of Economic Inequality*. 2006;4(1):107-122.

66. Cruz J, Torres J. ¿ De qué depende la satisfacción subjetiva de los colombianos? *Cuadernos de Economía*. 2006;25(45):131-154.

67. Frijters P, Haisken-DeNew JP, Shields MA. Money does matter! evidence from increasing real income and life satisfaction in east germany following reunification. *Am Econ Rev*. 2004;94(3):730-740.

68. Gerstenblüth M, Rossi M, Jewell RT. Health and happiness in uruguay. .

69. Frey BS, Stutzer A. Happiness, economy and institutions. *The Economic Journal*. 2000;110(466):918-938.

70. Lelkes O. Tasting freedom: Happiness, religion and economic transition. *Journal of Economic Behavior & Organization*. 2006;59(2):173-194.

71. Lucas RE, Clark AE, Georgellis Y, Diener E. Unemployment alters the set point for life satisfaction. *Psychological Science*. 2004;15(1):8.

72. Clark AE, Oswald AJ. Unhappiness and unemployment. *The Economic Journal*. 1994;104(424):648-659.

73. Brenes-Camacho G. Favourable changes in economic well-being and self-rated health among the elderly. *Soc Sci Med*. 2011.

74. The World Bank. Population projection tables by country and group. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTHEALTHNUTRITIONANDPOPULATION/EXTDATASTATISTICSHNP/EXTHNPSTATS/0,,contentMDK:21737699~menuPK:3385623~pagePK:64168445~piPK:64168309~theSitePK:3237118~isCURL:Y,00.html. Updated 2013. Accessed 04/12, 2013.

75. Klonowicz T. Discontented people: Reactivity and locus of control as determinants of subjective well-being. *European Journal of Personality*. 2001;15(1):29-47.

76. Minkov M. Predictors of differences in subjective well-being across 97 nations. *Cross-Cultural Research*. 2009;43(2):152-179.

77. Rabbitt P, Lunn M, Ibrahim S, Cobain M, McInnes L. Unhappiness, health and cognitive ability in old age. *Psychol Med*. 2008;38(2):229-236.

78. González-Quiñones JC, Restrepo-Chavarriaga G. Prevalencia de felicidad en ciclos vitales y relación con redes de apoyo en población colombiana. *Revista de Salud Pública*. 2010;12(2):228-238.

79. Lora E. *Beyond facts: Understanding quality of life.* Harvard Univ David Rockefeller; 2008.

80. Ferrer-i-Carbonell A, Frijters P. How important is methodology for the estimates of the determinants of happiness?*. *The Economic Journal*. 2004;114(497):641-659.

81. Rosero-Bixby L, Fernandez X, Dow WH. CRELES: Costa rican longevity and healthy aging (costa rica estudio de longevidad y envejecimiento saludable): Interviewer manual. *ICPSR*. 2005;26681.

82. Rosero-Bixby L, Fernandez X, Dow WH. CRELES: Costa rican longevity and healthy aging (costa rica estudio de longevidad y envejecimiento saludable): Sampling and methods - wave 2. *ICPSR*. 2008;26681.

83. Rosero-Bixby L, Fernandez X, Dow WH. CRELES: Costa rican longevity and healthy aging (costa rica estudio de longevidad y envejecimiento saludable): Weighting factors. *ICPSR*. 2005;26681.

84. Hardy SE, Allore H, Studenski SA. Missing data: A special challenge in aging research. *J Am Geriatr Soc*. 2009;57(4):722-729.

85. Diener E. Subjective well-being. *Psychol Bull*. 1984;95(3):542-575.

86. Clark AE, Oswald AJ. The curved relationship between subjective well-being and age. . 2006.

87. Zimmermann AC, Easterlin RA. Happily ever after? cohabitation, marriage, divorce, and happiness in germany. *Population and Development Review*. 2006;32(3):511-528.

88. Clark A, Lelkes O. Deliver us from evil: Religion as insurance. *Papers on Economics of Religion*. 2005;603:1-36.

89. Clark AE, Frijters P, Shields MA. Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic Literature*. 2008;46(1):95-144.

90. Kemm J. Towards an epidemiology of positive health. *Health Promot Internation*. 1993;8(2):129-134.

91. Almeida OP, Almeida SA. Short versions of the geriatric depression scale: A study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry*. 1999;14(10):858-865.

92. Rosero-Bixby L, Fernandez X, Dow WH. CRELES: Costa rican longevity and healthy aging study, 2005 ICPSR 26681 (costa rica estudio de

longevidad y envejecimiento saludable): Recoded variables. *ICSPR*. 2005(26681).

93. Spector WD, Fleishman JA. Combining activities of daily living with instrumental activities of daily living to measure functional disability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 1998;53(1):S46.

94. Metabolic syndrome - IDF criteria. http://en.wikipedia.org/wiki/Metabolic_syndrome#IDF. Accessed 07/20, 2012.

95. Juster RP, McEwen BS, Lupien SJ. Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews*. 2010;35(1):2-16.

96. Hasson D, Schwarz UVT, Lindfors P. Self-rated health and allostatic load in women working in two occupational sectors. *Journal of health psychology*. 2009;14(4):568-577.

97. Karlamangla AS, Singer BH, McEwen BS, Rowe JW, Seeman TE. Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. *J Clin Epidemiol*. 2002;55(7):696-710.

98. Angel RJ, Angel JL, Hill TD. Subjective control and health among mexican-origin elders in mexico and the united states: Structural considerations in comparative research. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2009;64(3):390.

99. Subramanian S, Delgado I, Jadue L, Vega J, Kawachi I. Income inequality and health: Multilevel analysis of chilean communities. *J Epidemiol Community Health*. 2003;57(11):844.

100. Instituto Nacional de Estadistica. Informe pobreza y desigualdad en uruguay, 2006. . 2007.

101. Instituto nacional de estadistica y censos de Costa Rica (INEC). Cuadro 4. limites de extrema pobreza y no extrema. 1987-2009. segun zona, . 2012.

102. Trujillo AJ, Portillo JE, Vernon JA. The impact of subsidized health insurance for the poor: Evaluating the colombian experience using propensity score matching. *International journal of health care finance and economics*. 2005;5(3):211-239.

103. Morris SS, Carletto C, Hoddinott J, Christiaensen LJM. Validity of rapid estimates of household wealth and income for health surveys in rural africa. *J Epidemiol Community Health*. 2000;54(5):381-387.

104. Royston P, White IR. Multiple imputation by chained equations (MICE): Implementation in stata. *Journal of Statistical Software*. 2011;45(4):1-20.

105. Stuart EA, Azur M, Frangakis C, Leaf P. Multiple imputation with large data sets: A case study of the children's mental health initiative. *Am J Epidemiol*. 2009;169(9):1133-1139.

106. Wooldridge JM. *Introductory econometrics: A modern approach.* South-Western Pub; 2009.

107. Roberto PN, Mitchell JM, Gaskin DJ. Plan choice and changes in access to care over time for SSI-eligible children with disabilities. *Journal Information*. 2005;42(2).

108. Williams R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal*. 2007;6(1):58-82.

109. Baetschmann G, Staub K, Winkelmann R. *Consistent estimation of the fixed effects ordered logit model.* Forschungsinstitut zur Zukunft der Arbeit GmbH; 2011.

110. Allison PD. Change scores as dependent variables in regression analysis. *Sociological methodology*. 1990;20(1):93-114.

111. Antonakis J, Bendahan S, Jacquart P, Lalive R. On making causal claims: A review and recommendations. *The Leadership Quarterly*. 2010;21(6):1086-1120.

112. StataCorp. Stata statistical software: Release 11. . 2009.

113. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173.

114. Lenze EJ, Rogers JC, Martire LM, et al. The association of late-life depression and anxiety with physical disability: A review of the literature and prospectus for future research. *The American Journal of Geriatric Psychiatry*. 2001;9(2):113-135.

115. Joseph S, Lewis CA. The Depression–Happiness scale: Reliability and validity of a bipolar self-report scale. *J Clin Psychol*. 1998;54(4):537-544.

116. di Cesare M, Guzman JM. Elderly well-being in latin america countries: Determinants and gender differences. .

117. Kunzmann U, Little TD, Smith J. Is age-related stability of subjective well-being a paradox? cross-sectional and longitudional evidence from the berlin aging study. *Psychol Aging*. 2000;15(3):511.

118. Blanchflower DG, Oswald AJ. Hypertension and happiness across nations. *J Health Econ*. 2008;27(2):218-233.

119. Mroczek DK, Kolarz CM. The effect of age on positive and negative affect: A developmental perspective on happiness. *J Pers Soc Psychol*. 1998;75(5):1333.

120. Taylor SE, Kemeny ME, Reed GM, Bower JE, Gruenewald TL. Psychological resources, positive illusions, and health. *Am Psychol*. 2000;55(1):99.

121. Guven C, Saloumidis R. Why is the world getting older? the influence of happiness on mortality. . 2009.

122. Seligman M. *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment.* Free Press; 2002.

123. Bago d'Uva T, Van Doorslaer E, Lindeboom M, O'Donnell O. Does reporting heterogeneity bias the measurement of health disparities? *Health Econ*. 2008;17(3):351-375.

124. Johnston DW, Propper C, Shields MA. Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. *J Health Econ*. 2009;28(3):540-552.

125. Bobak M, Pikhart H, Hertzman C, Rose R, Marmot M. Socioeconomic factors, perceived control and self-reported health in russia. A cross-sectional survey. *Soc Sci Med*. 1998;47(2):269-279.

126. Wilkinson W, Cato Institute. *In pursuit of happiness research: Is it reliable? what does it imply for policy?* Cato Institute; 2007.

127. Cheng H, Furnham A. Personality, self-esteem, and demographic predictions of happiness and depression. *Personality and Individual Differences*. 2003;34(6):921-942.

128. Soldz S, Vaillant GE. The big five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personality*. 1999;33(2):208-232.

129. Cheng ST. Age and subjective well-being revisited: A discrepancy perspective. *Psychol Aging*. 2004;19(3):409.

130. Sposito G, Diogo MJDE, Cintra FA, Neri AL, Guariento ME, De Sousa MLR. Relationship between subjective well-being and the functionality of elderly outpatients. *Revista Brasileira de Fisioterapia*. 2010;14(1):81-89.

131. Cid A, Ferrés D, Rossi M. Helping to unravel the dynamics of happiness among the elderly in the southern cone. *Revista de Ciencias Empresariales y Economía*. 2010(9):59-64.

132. Lyubomirsky S, Dickerhoof R, Boehm JK, Sheldon KM. Becoming happier takes both a will and a proper way: An experimental longitudinal intervention to boost well-being. *Emotion-APA*. 2011;11(2):391.

133. Vittersø J, Nilsen F. The conceptual and relational structure of subjective well-being, neuroticism, and extraversion: Once again, neuroticism is the important predictor of happiness. *Soc Indicators Res*. 2002;57(1):89-118.

134. Lu L. The relationship between subjective well-being and psychosocial variables in taiwan. *J Soc Psychol*. 1995;135(3):351-357.

135. Moreira-Almeida A, Lotufo Neto F, Koenig HG. Religiousness and mental health: A review. *Revista Brasileira de Psiquiatria*. 2006;28(3):242-250.

136. Barker DJP. Fetal origins of coronary heart disease. *BMJ*. 1995;311(6998):171.

137. Ozanne SE, Hales CN. Lifespan: Catch-up growth and obesity in male mice. *Nature*. 2004;427(6973):411-412.

138. Kuh D, Wadsworth MEJ. Physical health status at 36 years in a british national birth cohort. *Soc Sci Med*. 1993;37(7):905-916.

139. Taylor SE, Lerner JS, Sage RM, Lehman BJ, Seeman TE. Early environment, emotions, responses to stress, and health. *J Pers*. 2004;72(6):1365-1394.

140. Smith GD, Hart C, Blane D, Gillis C, Hawthorne V. Lifetime socioeconomic position and mortality: Prospective observational study. *BMJ*. 1997;314(7080):547.

141. Elstad JI. Childhood adversities and health variations among middle-aged men: A retrospective lifecourse study. *The European Journal of Public Health*. 2005;15(1):51-58.

142. Blanchflower DG, Oswald AJ. International happiness. *NEBR*. 2011.

143. Helmer C, Barberger-Gateau P, Letenneur L, Dartigues JF. Subjective health and mortality in french elderly women and men. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 1999;54(2):S84.

144. DeNeve KM, Cooper H. The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychol Bull*. 1998;124(2):197.

145. Murasko JE. A lifecourse study on education and health: The relationship between childhood psychosocial resources and outcomes in adolescence and young adulthood. *Soc Sci Res*. 2007;36(4):1348-1370.

146. Redelmeier DA, Katz J, Kahneman D. Memories of colonoscopy: A randomized trial. *Pain*. 2003;104(1):187-194.

147. Pelaez, Martha, Alberto Palloni, Cecilia Albala, Juan C. Alfonso, Roberto Ham-Chande, Anselm Hennis, Maria Lucia Lebrao, Esther Lesn-Diaz, Edith Pantelides, and Omar Prats. SABE- SURVEY ON HEALTH, WELL-BEING, AND AGING IN LATIN AMERICA AND THE CARIBBEAN, 2000 [computer file]. . 2000.

148. Health and retirement study - HRS 2010 - section B - demographics. . 2011.

149. Joia LC, Ruiz T, Donalisio MR. Life satisfaction among elderly population in the city of botucatu, southern brazil. *Revista de Saúde Pública*. 2007;41(1):131-138.

150. Trzcinski E, Holst E. Initial predictors of life satisfaction in early adulthood. *Schmollers Jahrbuch*. 2007;127(1):95-104.

151. Adam T, Lim SS, Mehta S, et al. Cost effectiveness analysis of strategies for maternal and neonatal health in developing countries. *BMJ*. 2005;331(7525):1107.

152. Tan-Torres Edejer T, Aikins M, Black R, Wolfson L, Hutubessy R, Evans DB. Cost effectiveness analysis of strategies for child health in developing countries. *BMJ*. 2005;331(7526):1177.

153. Easterlin RA. Explaining happiness. *Proc Natl Acad Sci U S A*. 2003;100(19):11176.

154. Reid MC, Papaleontiou M, Ong A, Breckman R, Wethington E, Pillemer K. Self-Management strategies to reduce pain and improve function among older adults in community settings: A review of the evidence. *Pain Medicine*. 2008;9(4):409-424.

155. Sung K. The effects of 16-week group exercise program on physical function and mental health of elderly korean women in long-term assisted living facility. *J Cardiovasc Nurs*. 2009;24(5):344.

156. Lam LCW, Lui VWC, Luk DNY, et al. Effectiveness of an individualized functional training program on affective disturbances and functional skills in mild and moderate dementia—a randomized control trial. *Int J Geriatr Psychiatry*. 2009;25(2):133-141.

157. Chang JT, Morton SC, Rubenstein LZ, et al. Interventions for the prevention of falls in older adults: Systematic review and meta-analysis of randomised clinical trials. *BMJ*. 2004;328(7441):680.

158. Alibhai SMH, Greenwood C, Payette H. An approach to the management of unintentional weight loss in elderly people. *Can Med Assoc J*. 2005;172(6):773-780.

159. Conn VS, Hafdahl AR, Brown SA, Brown LM. Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. *Patient Educ Couns*. 2008;70(2):157.

160. Forsman AK, Nordmyr J, Wahlbeck K. Psychosocial interventions for the promotion of mental health and the prevention of depression among older adults. *Health Promot Internation*. 2011;26(suppl 1):i85-i107.

161. Forte A, Hill M, Pazder R, Feudtner C. Bereavement care interventions: A systematic review. *BMC Palliative Care*. 2004;3(1):3.

162. Lapierre S, Erlangsen A, Waern M, et al. A systematic review of elderly suicide prevention programs. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*. 2011;32(2):88-98.

163. Coberley C, Rula EY, Pope JE. Effectiveness of health and wellness initiatives for seniors. *Population Health Management*. 2011;14(S1):45-50.

164. McCusker J, Verdon J. Do geriatric interventions reduce emergency department visits? A systematic review. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2006;61(1):53-62.

165. Markle-Reid M, Browne G, Weir R, Gafni A, Roberts J, Henderson SR. The effectiveness and efficiency of home-based nursing health promotion for older people: A review of the literature. *Medical Care Research and Review*. 2006;63(5):531-569.

166. Bianca B, Juliette P, van Deelen Bob HR. A randomised clinical trial on a comprehensive geriatric assessment and intensive home follow-up after hospital discharge: The transitional care bridge. *BMC Health Services Research*. ;10.

167. Boren SA, Fitzner KA, Panhalkar PS, Specker JE. Costs and benefits associated with diabetes education A review of the literature. *Diabetes Educ*. 2009;35(1):72-96.

168. Grabowski DC. The cost-effectiveness of noninstitutional long-term care services: Review and synthesis of the most recent evidence. *Medical Care Research and Review*. 2006;63(1):3-28.

169. Robertson MC, Gardner MM, Devlin N, McGee R, Campbell AJ. Effectiveness and economic evaluation of a nurse delivered home exercise programme to prevent falls. 2: Controlled trial in multiple centres. *BMJ*. 2001;322(7288):701.

170. Huang ES, Zhang Q, Brown SES, Drum ML, Meltzer DO, Chin MH. The Cost-Effectiveness of improving diabetes care in US federally qualified community health centers. *Health Serv Res*. 2007;42(6p1):2174-2193.

171. Savelkoul M, de Witte L, Post M. Stimulating active coping in patients with rheumatic diseases: A systematic review of controlled group intervention studies. *Patient Educ Couns*. 2003;50(2):133-143.

172. Gross national happiness. http://www.grossnationalhappiness.com/articles/. Accessed 02/02, 2013.

173. Fernández X, Robles A. I informe estado de situación de la persona adulta mayor en costa rica. *UCR, CONAPAM.San José, Costa Rica*. 2008.

174. Costa Rican Social Security (CCSS). Memoria X aniversario ciudadano de oro. .

175. Ross A.
 **
Carne de oro de CCSS trae descuentos a ancianos** http://www.nacion.com/2012-05-17/ElPais/carne-de-oro-de-

ccss-trae-descuentos-a--ancianos.aspx. Updated 2012. Accessed 011/19, 2012.

176. Bucheli M, Rossi M. El grado de conformidad con la vida: Evidencia para las mujeres del gran montevideo. *Documentos de Trabajo (working papers)*. 2003.

177. Bjørnskov C. The happy few: Cross-country evidence on social capital and life satisfaction. *Kyklos*. 2003;56(1):3-16.

178. Ritchey LH, Ritchey PN, Dietz BE. Clarifying the measurement of activity. *Act Adapt Aging*. 2001;26(1):1-21.

179. Burt DB, Zembar MJ, Niederehe G. Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychol Bull*. 1995;117(2):285.

180. Hernandez LC, Slon P. TIPOLOGÍA DE LA POBREZA SEGÚN LA TEORÍA DEL CICLO DE LA VIDA. .

181. Poirier DJ. Partial observability in bivariate probit models. *J Econ*. 1980;12(2):209-217.

182. Imbens GM, Wooldridge JM. Recent developments in the econometrics of program evaluation. *IZA Discussion Papers*. 2008;3640.

183. Chiburis RC, Das J, Lokshin M. A practical comparison of the bivariate probit and linear IV estimators. *Economics Letters*. 2012.

184. Nichols A. Causal inference for binary regression. . 2011;14:2011.

185. Imbens GW, Lemieux T. Regression discontinuity designs: A guide to practice. *J Econ*. 2008;142(2):615-635.

186. Nichols A. Rd 2.0: Revised stata module for regression discontinuity estimation. . 2011.

187. Görg H, Strobl E. The effect of R&D subsidies on private R&D. *Economica*. 2007;74(294):215-234.

188. Norton EC, Wang H, Ai C. Computing interaction effects and standard errors in logit and probit models. *Stata Journal*. 2004;4:154-167.

189. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127:757-763.

190. Nichols A. Causal inference with observational data. *Stata Journal*. 2007;7(4):507.

191. Leuven E, Sianesi B. PSMATCH2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing"
. . ;version 4.0.6.

192. Becker SO, Ichino A. Estimation of average treatment effects based on propensity scores. *The stata journal*. 2002;2(4):358-377.

193. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*. 1985;39(1):33-38.

194. Schafer JL, Kang J. Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychol Methods*. 2008;13(4):279.

195. Wang M. Profiling retirees in the retirement transition and adjustment process: Examining the longitudinal change patterns of retirees' psychological well-being. *J Appl Psychol*. 2007;92(2):455.

196. Imbens G, Kalyanaraman K. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*. 2012;79(3):933-959.

197. Frey BS, Stutzer A. Testing theories of happiness. *Institute for Empirical Research in Economics.Working Paper Series*. 2003;147.

198. Di Tella R, MacCulloch RJ, Oswald AJ. Preferences over inflation and unemployment: Evidence from surveys of happiness. *Am Econ Rev*. 2001:335-341.

199. Camfield L, Skevington SM. On subjective well-being and quality of life. *Journal of health psychology*. 2008;13(6):764.

200. Headey B, Wearing AJ. *Understanding happiness: A theory of subjective well-being.* Longman Cheshire; 1992.

201. Easterlin RA. A puzzle for adaptive theory. *Journal of Economic Behavior & Organization*. 2005;56(4):513-521.

202. Jasti S, Dudley WN, Goldwater E. SAS macros for testing statistical mediation in data with binary mediators or outcomes. *Nurs Res*. 2008;57(2):118.

203. Lyubomirsky S, Lepper HS. A measure of subjective happiness: Preliminary reliability and construct validation. *Soc Indicators Res*. 1999;46(2):137-155.

204. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess*. 1985;49(1):71-75.

205. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *J Pers Soc Psychol*. 1988;54(6):1063.

206. Lane RE. *The loss of happiness in market democracies.* Yale University Press; 2001.

207. Frey BS, Stutzer A. *Happiness and economics: How the economy and institutions affect human well-being.* Princeton University Press; 2001.

208. Meyer C, Kirby J. **Is GDP the right measure of wealth and well-being?**. http://blogs.hbr.org/hbr/meyer-kirby/2011/03/wealth-and-well-being-the-lega.html?cm_sp=blog_flyout-_-hbrmeyer-kirby-_-wealth_and_well-being_the_lega. Updated 2011. Accessed 01/30, 2013.

209. Cummins RA, Lau AALD, Mellor D, Stokes MA. Encouraging governments to enhance the happiness of their nation: Step 1: Understand subjective wellbeing. *Soc Indicators Res*. 2009;91(1):23-36.

210. Fujita F, Diener E. Life satisfaction set point: Stability and change. *J Pers Soc Psychol*. 2005;88(1):158.

211. Gamester N, Lovo S, Masino S, Omic E. The 2012 legatum prosperity index: Methodology and technical appendix. . 2012.

212. Gruber J, Mullainathan S. Do cigarette taxes make smokers happier? *Do cigarette taxes make smokers happier?*. 2002.

213. Andreasen NC, Scheftner W, Reich T, Hirschfeld R, Endicott J, Keller MB. The validation of the concept of endogenous depression: A family study approach. *Arch Gen Psychiatry*. 1986;43(3):246.

214. Vikander N. *Kahnemanrs objective happiness and senrs capabilities: A critical comparison*. 2007.

215. OECD - your better life index. http://www.oecdbetterlifeindex.org/. Accessed 07/09, 2013.

216. Kittiprapas S, Sawangfa O, Fisher C, Powdthavee N, Nitnitiphrut K. Happiness: New paradigm, measurement, and policy implications. . 2007.

CURRICULUM VITAE

ALEXO ESPERATO-MARTÍNEZ

3100 18th St. NW
Washington DC 20010 – USA

aesperato@yahoo.com
Tlf: +202 468 8262

PROFESSIONAL EXPERIENCE

INTERNATIONAL CONSULTING PROJECTS

Healthcare access expansion for vulnerable populations in Panama *(Dec. 2012 – Jan 2013)*

- Developed a methodological proposal for the project auditor (Gesaworld, a consulting company). The project is a joint initiative of the World Bank and the Inter-American Development Bank

Micronutrient guideline development at World Health Organization *(Dec 2011 – Feb 2012)*

- Developed a methodological proposal for the evaluation of the guideline development process at WHO-Geneva. The evaluation is scheduled to proceed onsite in March-August of 2013

Workers' compensation strategy at American International Group (AIG) *(Mar.– Dec. 2011)*

- Researched the main determinants of workers' compensation expenditures in 2000-2012
- Researched and assisted in the coordination of a task force to examine the main causes of workers' compensation claims in 2012-2020

Malaria reduction strategy in Latin America and the Caribbean *(Jul.– Oct. 2009)*

- Performed a desk review and evaluated this project onsite (Panama) for the Pan American Health Organization/World Health Organization

PAN AMERICAN HEALTH ORGANIZATION (PAHO/WHO), Washington, DC

Consultant at Quality of Health Services and Patient Safety *(Jan. 2007 – Jun. 2009)*

- Designed and coordinated quality of health services projects in Latin America and the Caribbean that collaborated with over 10 countries in the region
- Projects encompassed various areas, including healthcare-related infections, adverse events associated with surgery, civil society, and research
- Main tasks included project design, political/resource mobilization, institutional coordination, and training of government officials and civil society representatives

Consultant at Regional Forum for Public Health in the Americas *(Jan. 2006 – Dec. 2006)*

- Executed a systematic review and taxonomy of existing virtual policy networks and think tanks
- Elaborated executive summaries and proposals to PAHO's Executive Management

- Performed research studies and technical reports in public policy and knowledge-management areas

BARCELONA PUBLIC HEALTH AGENCY (ASPB), Barcelona, Spain **(***Sep. 2005 – Dec. 2006)*
Research Associate

- Executed a systematic review of the economic evaluations of health prevention interventions in Spain

MÉDICUS MUNDI CATALUNYA (MMC), Barcelona, Spain *(March– Aug. 2005)*
Technical Officer

- Performed financial control and reporting to main project donors (European Union, AECI)
- Analyzed the socioeconomic context, health, and international aid policies. The resulting report was a main technical input for Medicus Mundi's strategy in Mozambique.

AGILENT TECHNOLOGIES, Barcelona, Spain (*Jul. 2002 – Feb. 2005)*
Accounting Expert

- Responsible for the Reconciliation and investigation of Brazilian accounts receivables

DISTRIBUIDORA LUNAR, Vitória (ES), Brazil *(Sep. 2001 – Apr. 2002)*
General Manager Staff

- Supported general management in the enterprise strategy

ADEXA, INC. Munich, Germany; and Brussels, Belgium *(Jul. 1999 – Aug. 2001)*
Business Development Manager

- Coordinated business development activities for the European markets

ASIAN-PACIFIC ECONOMIC COOPERATION (APEC), Portland, OR *(Sep. 1997 – Dec.1997)*
Research Intern

LANGUAGES

Spanish: Native. **English, German, French, Italian, Portuguese:** Fluent

SOFTWARE

Financial Modeling in Lotus 1-2-3. Statistical analysis in **Shazam 7.0, SPSS, Stata** and **MPLUS.** Finance and administration in **SAP R/3/ Oracle Applications.** Database Management in **ACT!** Decision modeling in **Treeage Pro.** Qualitative Data Analysis in **Atlas Ti.**

EDUCATION

JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH, Baltimore, MD
PhD in Health Services Research. Aug. 2009 – Feb 2014 (exp.)

UNIVERSIDAD POMPEU FABRA - THE JOHNS HOPKINS UNIVERSITY, Barcelona, Spain
Master in Public and Social Policy
Sept 2003 - Dec. 2005

ESCI-UNIVERSIDAD POMPEU FABRA, Barcelona, Spain
Master of International Business
Aug. 2002- Jul. 2003

LEWIS & CLARK COLLEGE, Portland OR
Aug. 1994 - May 1998
B.A. Economics

PUBLICATIONS

Esperato, A. Bishai, D. Projecting the health and economic impact of road safety initiatives: a case study of a multi-country project. Traffic Inj Prev. 2012;13 Suppl 1:82-9.

Esperato, A. García-Altés, A. Health promotion: a profitable investment? Economic efficiency of preventive interventions in Spain. Gac Sanit. 2007 Mar-Apr;21(2):150-61.

Espíritu, Nora. Notification of adverse events in a national hospital in Lima. Revista Española de Calidad Asistencial, Dec 2007. [Acknowledged]

Gesaworld. Evaluaciones de políticas y programas de salud: conceptos, metodologías y experiencias. Barcelona, 2010. [Acknowledged]

AWARDS AND DISTINCTIONS

Johns Hopkins Bloomberg School of Public Health *(2009-2013)*
Sommer Scholar