KU LEUVEN

DEPARTMENT OF ECONOMICS

# Program evaluation and ethnic differences: the Pension 65 program in Peru

Koen DECANCQ, Javier OLIVERA and Erik SCHOKKAERT

FACULTY OF ECONOMICS AND BUSINESS

DISCUSSION PAPER SERIES DPS18.14

NOVEMBER 2018

# Program evaluation and ethnic differences: the Pension 65 program in Peru*

Koen Decancq† Javier Olivera‡ Erik Schokkaert§

November 21, 2018

**Abstract**

We show that the introduction of a non-contributory pension program (Pension 65) in Peru had remarkably different effects for its three main ethnic groups, i.e., Mestizo, Quechua, and Aymara. The Aymara beneficiaries of the program have experienced larger increases in health and life satisfaction compared to other Peruvians. Using a panel life satisfaction regression, we find evidence for preference heterogeneity between the Aymara and the other ethnic groups that is consistent with the observed differences. Finally, we turn to the question of how the pension program can be evaluated in a robust manner while respecting the preference heterogeneity between the ethnic groups. We propose the natural criterion that a program benefits a recipient if she is lifted to a higher indifference curve. We show that the pension program was beneficial for all groups, but that more Aymara beneficiaries were lifted to a higher indifference curve compared to other Peruvians. Our proposed criterion can be useful to evaluate programs in all cases where preference differences matter.

**Keywords**: Program Evaluation, Ethnicity, Preferences, Pensions, Peru.

**JEL classification**: O12, D12, I15, I38.

*We thank Sam Cosaert, Marc Craps, Shaun Da Costa, Tom Van Ourti, and participants in presentations in Leuven, Antwerp, Brussels, Luxembourg, and Tampere for helpful discussions and comments.

†Centre for Social Policy (University of Antwerp), Centre for Philosophy of Natural and Social Science (London School of Economics), CORE (Université catholique de Louvain-la-Neuve), and Department of Economics (KULeuven).

‡LISER (Luxembourg Institute of Socio-Economic Research) and Department of Economics (Pontificia Universidad Catolica del Peru).

§Department of Economics (KULeuven) and CORE (Université catholique de Louvain-la-Neuve).

# 1 Introduction

In recent decades there has been a growing attention in policy circles for the specific position of indigenous people in the process of development (see, e.g., the "Indigenous and Tribal Peoples Convention" of the International Labor Organization or the United Nations Declaration 61/295 on the Rights of Indigenous Peoples).[1] There are indeed wide socio-economic gaps between indigenous and non-indigenous citizens. Moreover, indigenous people may have different values and preferences, and it has been argued that these differences should be respected when evaluating their overall situation. Yet, the latter principle seems to have had only a minimal impact on the concrete practice of making program evaluations.

In this paper, we evaluate the effects of the introduction of the Pension 65 program, a non-contributory pension program in Peru. Even though the Pension 65 program is in the first place an income transfer program, its beneficiaries can also receive care in public health facilities at no cost and are eligible for the "Integral Health Insurance Plan". Like other Latin American countries, Peru is characterized by a large socio-economic gap between indigenous and non-indigenous people (World Bank, 2015). We focus on the effects of the policy on the two largest indigenous groups, the Quechua and the Aymara, in comparison with the group of citizens who have a mixed, Mestizo, identity.

To quantify the effects of the Pension 65 program, we use panel data from the Survey of Health and Wellbeing of the Elderly (ESBAM) that was carried out before and after the introduction of the Pension 65 program (in 2012 and 2015).[2] We use a panel regression difference-in-differences estimator exploiting the discontinuity at the eligibility threshold. We compare the beneficiaries of the Pension 65 program to a control group of non-beneficiaries who are classified as poor, but not as extremely poor and, hence, were not eligible. Remarkably, we find that the Aymara beneficiaries of the Pension 65 program have experienced

[1] Art. 2(1c) of *The Indigenous and Tribal Peoples Convention 169* of the International Labor Organization (1989) reads "[Governments should take measures for] assisting the members of the peoples concerned to eliminate socio-economic gaps that may exist between indigenous and other members of the national community, in a manner compatible with their aspirations and ways of life." The *United Nations Declaration 61/295 on the Rights of Indigenous Peoples* (2007) includes Art. 21(2): "States shall take effective measures and, where appropriate, special measures to ensure continuing improvement of their [indigenous peoples] economic and social conditions", and Art. 33(1): "Indigenous peoples have the right to determine their own identity or membership in accordance with their customs and traditions."

[2] The ESBAM data have been collected by the National Institute of Statistics and Informatics of Peru (INEI) with the specific purpose to study the impact of the Pension 65 program.

larger increases in health and life satisfaction compared to other Peruvians.

Although it is common practice in the literature to evaluate the effects of policies dimension by dimension, it is useful to integrate these effects in an overall measure of individual well-being. This integration allows us to make an overall assessment of the program and to shed light on the phenomenon of cumulative deprivation, which arises when the same individuals are deprived in different dimensions of life (Ferreira and Lugo, 2013). One possible candidate for such a measure of individual well-being is self-reported subjective life satisfaction. We will propose an alternative measure that is solely based on the ordinal preferences of the Peruvians in the different ethnic groups, capturing what they themselves consider to be a "good life". We say that a program benefits a recipient if she is lifted to a higher indifference curve according to her own preferences. This criterion makes use only of ordinal intra-personal well-being comparisons and does not involve any debatable normative choices that are needed to cardinalize utility in an interpersonally comparable way. It is consistent with various preference-based well-being measures based on quantity metrics or money metrics (see Deaton and Muellbauer (1980, pp. 179-182) for a survey).

Following Decancq et al. (2015a; 2015b; 2017), we obtain information about the differences in preferences between the ethnic groups from a life satisfaction regression (see also Clark and Oswald (2002); Van Praag and Baarsma (2005)). We find that the Aymara assign a relatively large weight to health in their conception of a good life. This finding is consistent with the observed heterogeneity in outcomes of the Pension 65 program. It is also in line with some anthropological work on the features of what could be a specific Andean view on the good life (see Smith (2006) on the Quechua, and Calestani (2009) on the Aymara). Calestani (2009) describes the Aymara view on the good life in terms of the basic concepts of *suma qamaña* ("living well together", i.e., in harmony with society) and *suma jakaña* (living in harmony with oneself). Even though these idealized constructions do not exist in reality (Artaraz and Calestani, 2015), they function as a general frame of reference to evaluate the real-world situation. From this perspective, it is to be expected that a non-material life dimension such as health becomes a relatively important component of the good life.

Our analysis is related to several recent contributions. Copestake et al. (2009) give an overview of the mismatch between subjective well-being and simple monetary measures of poverty in Peru, but they do not use the life satisfaction regression to derive information about marginal rates of substitution between

different life dimensions. Bernal et al. (2017) and Neelsen and O'Donnell (2017) have investigated the impact of the extension of health insurance in Peru (the "Integral Health Insurance Plan") on health care access and health care expenditures without including other dimensions of well-being. Olivera and Tournier (2016) use the baseline wave of the ESBAM data to analyze the well-being of the poor elderly in Peru. They use a multidimensional counting approach and do not focus on the impact of the Pension 65 program. Closest to our paper is the evaluation of the Pension 65 program by Bando et al. (2016). They look at the intention-to-treat effect, rather than at the treatment effect and use a different econometric strategy to identify the effects of the Pension 65 program. Bando et al. (2016) analyze the impact of the program on different life dimensions separately without looking at an overall well-being measure. None of the papers on Peru have explicitly analyzed ethnic differences. For Bolivia, however, Canavire Bacarreza et al. (2017) show that an unconditional cash transfer program has a larger effect on educational expenditures for indigenous than for non-indigenous recipients. Van de gaer et al. (2013) find stronger effects of Mexico's Oportunidades Program on the health opportunities of children from indigenous backgrounds than on the health opportunities of children from non-indigenous backgrounds. Kant et al. (2014) investigate ethnic differences in Canada on the basis of subjective well-being information, but they do not construct a preference-based well being measure.

The paper is organized as follows. In Section 2, we describe the background about the ethnic differences in Peru, the Pension 65 program, and the data set. In Section 3, we show that the Pension 65 program has different effects on the three ethnic groups and most strikingly so for the Aymara. In Section 4, we investigate the latter finding in more detail and we will argue that the estimated preference differences are consistent with our findings. The resulting normative issues and our preference-based criterion for policy evaluation are discussed in Section 5. We find that taking into account ethnic differences in the conception of a good life is relevant for the evaluation of the Pension 65 program. Section 6 concludes.

# 2 Background and data

## 2.1 Measuring ethnicity in Peru

The complex interplay between ethnicity, exclusion, segregation, and race makes it difficult to define and measure ethnicity in quantitative studies.[3] In Peru, ethnicity has functioned less as a basis for political organization than in other Latin American countries, although there seems to be a revival since the turn of the century (Sulmont, 2011). Still, as elsewhere, being indigenous in Peru is associated with less education, less income, more poverty, less power, and more negative stigmatization.

In general, there are two main approaches to measure ethnicity. A first approach is to use external observable characteristics such as mother tongue, physical characteristics, and place of origin.[4] Before 2017, Peru used mother tongue to demarcate between ethnic groups in its census (World Bank, 2015). The most common languages learned during childhood are Spanish (81%), Quechua (17%), and Aymara (2%), though there are other languages spoken in the Peruvian jungle and in the low areas between the Andes and the jungle. Mother tongue has been the traditional criterion to distinguish ethnic groups in multilingual countries. It is a problematic criterion, however: many people with indigenous origin speak Spanish and many indigenous people of different ethnicity speak Quechua (Paredes, 2007).

An alternative approach based on self-identification with a particular ethnic group has become dominant in the recent literature.[5] The 2007 Peruvian census sets the number of indigenous-language speakers at 4.4 million, whereas projections based on self-identification in household surveys yield an estimate of about 9.7 million indigenous persons (World Bank, 2015).[6] This is not to say that the self-identification approach is without problems. In a context of discrimination, individuals may be reluctant to reveal their indigenous ethnic background and may prefer to report a more neutral background such as being Mestizo (mixed

[3] The specific problems of defining ethnicity in Peru are further discussed in Paredes (2007); Sulmont (2011); Moreno (2014); Pasquier-Doumer and Risso Brandon (2015).

[4] The 19th century description of the "Aymara Indians" by Forbes (1870) compares at length the bodily measurements of the Aymara to people with a European or African background, for instance.

[5] ILO Convention 169 (1989) supported this idea in its Article 1(2): "Self-identification as indigenous or tribal shall be regarded as a fundamental criterion for determining the groups to which the provisions of this Convention apply."

[6] The 2017 census in Peru has used a self-identification approach.

background) to avoid the stigma of being non-white (Ñopo et al., 2004; Pasquier-Doumer and Risso Brandon, 2015). As a consequence, ethnic identification can be influenced by the relative numbers of indigenous people in the localities (Sulmont, 2011; Moreno, 2014). Indigenous individuals living in urban areas, where stigmatization is more salient, may prefer to identify as Mestizo. In contrast, individuals living in the Andes and rural areas of Peru with a higher concentration of indigenous people may have less resistance identifying as indigenous. Rapid urbanization in Peru makes ethnic identity particularly fluid. This is particularly true for the Quechua, who have moved more often from rural areas to the cities, than for the Aymara, who remain mainly concentrated in the rural highlands of Southern Peru.

In the analysis of this paper, we follow the dominant approach in the literature and use self-identification to measure ethnicity.[7] Our aim is to identify groups with a strong identity, i.e., groups who share a specific view on what is important in life. It seems reasonable to assume that the individuals who have overcome stigmatization and self-identify as being indigenous, are the ones who show a stronger and more entrenched identity (see also Sulmont, 2011).

## 2.2 The Pension 65 program

The non-contributory pension program Pension 65 has been introduced in October 2011 and is administered by the Ministry of Development and Social Inclusion of Peru (MIDIS). The program was rolled out between 2012 and 2014, and by 2015 it had reached about 500,000 Peruvians. The Pension 65 program is the second largest social program in Peru, behind the conditional cash transfer program "Juntos" (Bando et al., 2016).

The program has two components. First, beneficiaries receive 250 Peruvian soles (about US$76) every two months. This transfer amounts to about 18% of the total expenditures of the targeted group before the intervention. In addition, beneficiaries are eligible for the "Integral Health Insurance Plan" and they can receive care in public health facilities at no cost. Only individuals aged 65 or over, who are neither affiliated to any pension system nor already receiving a pension, and who are living in a household classified as "extremely poor" by

[7] There are exceptions to this trend. Pasquier-Doumer and Risso Brandon (2015) prefer to use language as the criterion, also because they focus on children. Canavire Bacarreza et al. (2017) combine self-identification and language to distinguish indigenous and non-indigenous groups in Bolivia.

the official targeting system SISFOH (Sistema de Focalización de Hogares), are eligible to the Pension 65 program.

In the SISFOH targeting system, every household obtains a score, which is unknown to them. A household's SISFOH score is a weighted average of a number of variables related to the material and socio-economic conditions of the household and its members.[8] Based on a comparison of their SISFOH score with region-specific poverty thresholds, households are classified into three categories: extremely poor, non-extremely poor, and non-poor. Only the first category is eligible for the Pension 65 program. The SISFOH scores were determined before the thresholds were established, which avoids the possibility of manipulation (see Camacho and Conover, 2011).

## 2.3 Data

We utilize data from the Survey of Health and Wellbeing of the Elderly (Encuesta de Salud y Bienestar del Adulto Mayor, or ESBAM) that was carried out in 2012 and 2015.[9] This panel data set is specifically intended to study the impact of the Pension 65 program on the elderly poor population. The data set is composed of detailed questionnaires for the individuals aged between 65 and 80 that include individualized information about their living standards, consumption, demographics, well-being, beliefs, time use, nutrition, and subjective and objective health variables, among others.

The data was gathered in 12 out of 24 departments of Peru in which the SISFOH registers have been updated at the moment of the sampling (unshaded districts in Figure 1 belong to departments that are outside the sampling frame). The sampling frame of ESBAM includes households with at least one member aged between 65 and 80 and having a SISFOH score above or below 0.3 standard deviations of the SISFOH threshold for extreme poverty. The goal of this design

[8] The SISFOH score includes information about the type of fuel used for cooking; the access to water, sewerage, electricity and telephone; the type of walls, roof and floor; the education of the head of household and the maximum level of education at home, health insurance, assets, and the extent of overcrowding.

[9] The survey has been administered by the National Institute of Statistics and Informatics of Peru (INEI). The questions appearing in ESBAM are inspired by leading old age surveys such as the Health and Retirement Study (HRS) or the Survey of Health, Ageing and Retirement in Europe (SHARE). The information is collected by means of face-to-face interviews by INEI's interviewers, while some bio-markers such as blood samples, arterial pressure and anthropometric measurements are collected by medical technicians during the fieldwork. The interviews for the baseline of 2012 were carried out in November and December, and the 2015 follow-up was carried out between July and October.

was to obtain households located sufficiently close to the eligibility threshold for the Pension 65 program, such that they would be similar in all relevant dimensions except for the eligibility threshold itself. We say that individuals living in households located below or above the eligibility threshold fall into the treatment and control group, respectively.[10]

The sampling procedure of the data set is probabilistic, independent in each department, and stratified by rural and urban areas. It has been carried out in two steps. In the first step, the primary sampling units in urban areas are the census units, in rural areas they coincide with villages with at least four households who are living in poverty and with at least one member older than 65. The selection probability of a primary sampling unit is proportional to the total number of households in the primary sampling unit. In the second step, four households were randomly drawn from each primary sampling unit. The initial sample size in the baseline survey of 2012 consists of 4,242 individuals (in 3,194 households) and there are 3,847 individuals (in 2,967 households) in the follow-up wave of 2015.

We restrict the sample to the respondents who are eligible for the Pension 65 program (apart from the threshold), have no missing data and are present in both waves (more details can be found in the Appendix). We focus on the respondents who self-identify as belonging to one of the three largest ethnic groups in Peru: Mestizo, Quechua, and Aymara. Though most respondents identify themselves with the same ethnic group in both waves, some change group across waves. The Aymara tend to self-identify most consistently across both waves. Only 7% of the respondents who self-identify as Aymara in the 2012 wave have changed to another ethnic group in the follow-up. Respondents who self-identify as Quechua and Mestizo in the baseline have changed ethnic group more often: 21% and 27% respectively. Given our interest in studying well-being and preferences by ethnic groups, we restrict our sample to those respondents who answered consistently the same ethnic group in both waves.[11] Our final sample is composed of 1,968 respondents who are observed in both waves, from which 1,233 (63%), 573 (29%) and 162 (8%) respondents are Mes-

[10] Unlike Bando et al. (2016), we were not granted access to the SISFOH scores of the respondents. This prevents us from removing respondents from the sample with SISFOH scores that are outside the admissible range of 0.3 standard deviations above or below the threshold, as Bando et al. (2016, fn. 5) do in their main specification. However, the authors report that the removal of these respondents is "not likely to affect the results".

[11] We also implemented a robustness check with a sample including the cases of respondents identifying with one of the above three ethnic categories in one wave and answering "don't know/no answer" in the other wave. This did not change our findings.

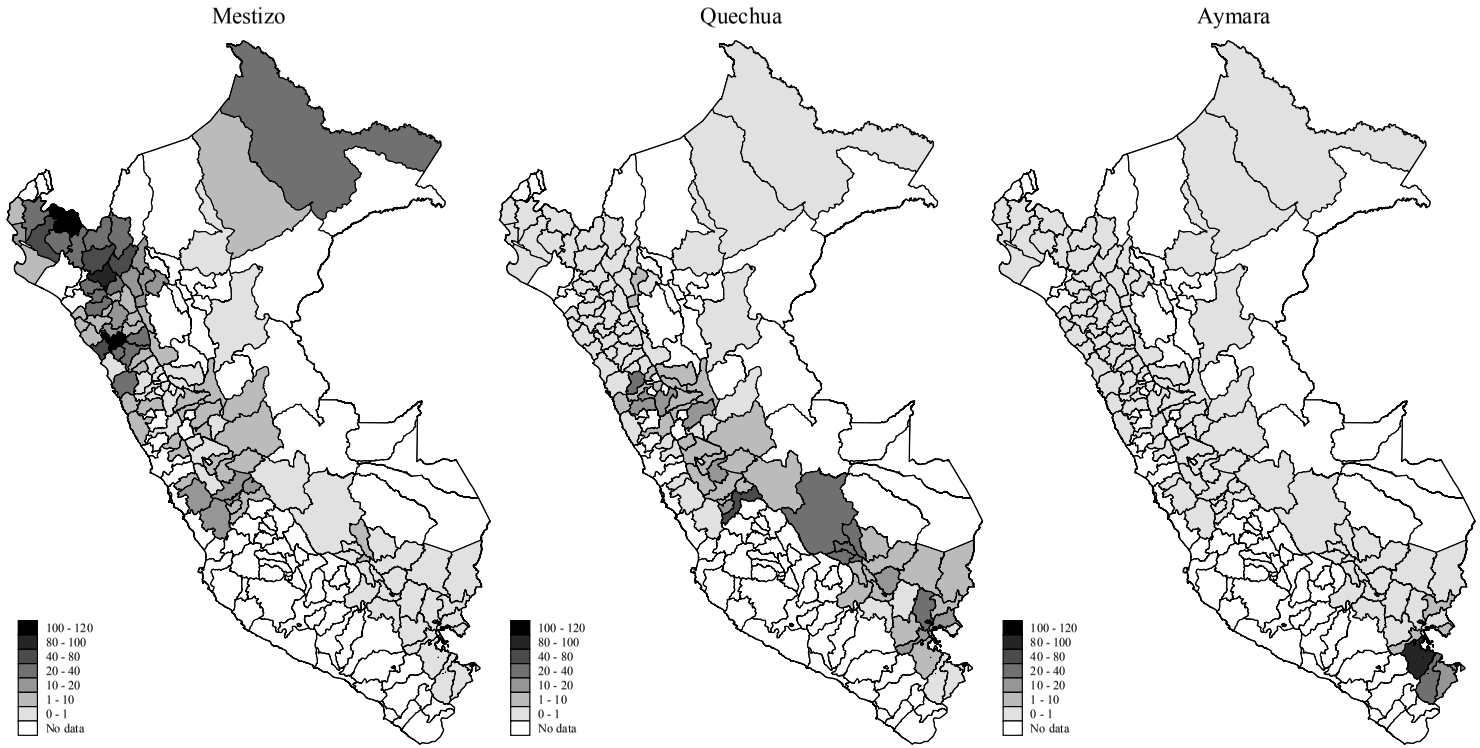


Figure 1: Ethnicity by region

tizo, Quechua and Aymara, respectively (see Table 1). Clearly, the final sample of 1,968 respondents is not representative for the entire Peruvian population, nor for the subpopulation targeted by ESBAM. This is not problematic for our purposes, since we want to focus primarily on differences between the ethnic groups.

| | Mestizo | Quechua | Aymara | total |
|---|---|---|---|---|
| control | 635 | 334 | 98 | 1,067 |
| treatment | 598 | 239 | 64 | 901 |
| total | 1,233 | 573 | 162 | 1,968 |

Table 1: Control and treatment by ethnicity

The respondents who identify as Aymara live more in rural areas (the urbanization rate of the Aymara sub sample is below 10%, whereas it is around 40% for the other groups). As can be seen in Figure 1, the Aymara live concentrated in the Altiplano region of Southern Peru, close to the Bolivian border and the Titicaca lake. Moreover, they live on average on a higher altitude (3900 km above sea-level), compared to the Quechua (3250 km) and Mestizo (1750 km). Given the concentration of the Aymara in a few specific districts, it is difficult to disentangle the effect of ethnicity from locality for this group. Where needed, we will control for regional effects and altitude of the district in which the respondent lives.

# 3 The impact of the Pension 65 program

## 3.1 A first glance at some key variables

Let us start by looking at the five important aspects of life that will play a central role in our analysis. Table 2 presents averages for 2012 and 2015 broken down by ethnic group and treatment status.

Information about *household expenditures* is collected for eight types of goods, by source of acquisition (buying, gift, and self-consumption).[12] Expenditures on

[12]The eight expenditure categories are food, clothes, utilities, durables, health, transport, leisure and other. The methodology to collect information on expenditures in the ESBAM data set follows closely that of the ENAHO survey, which is used by the Peruvian National Institute of Statistics to estimate official poverty rates.

| | | total | Mestizo | | Quechua | | Aymara | |
|---|---|---|---|---|---|---|---|---|
| | | | C | T | C | T | C | T |
| 2012 | expenditures (in logarithm) | 5.76 | 5.87 | 5.74 | 5.79 | 5.62 | 5.52 | 5.53 |
| | health (between 0 and 100) | 50.08 | 51.97 | 50.31 | 49.32 | 47.23 | 49.35 | 45.07 |
| | ADL (between 0 and 100) | 79.85 | 82.46 | 81.70 | 77.89 | 74.61 | 74.00 | 75.48 |
| | respect (between 0 and 100) | 84.13 | 87.56 | 87.73 | 80.89 | 78.28 | 71.18 | 75.00 |
| | SWB (between 0 and 100) | 53.40 | 56.69 | 56.79 | 48.82 | 49.16 | 44.23 | 42.63 |
| 2015 | expenditures (in logarithm) | 5.79 | 5.89 | 5.82 | 5.82 | 5.79 | 5.14 | 5.47 |
| | health (between 0 and 100) | 48.70 | 49.84 | 49.31 | 45.45 | 45.54 | 51.30 | 56.39 |
| | ADL (between 0 and 100) | 75.53 | 78.00 | 79.64 | 69.13 | 66.23 | 75.39 | 80.75 |
| | respect (between 0 and 100) | 82.73 | 85.82 | 85.30 | 77.16 | 76.61 | 78.54 | 86.51 |
| | SWB (between 0 and 100) | 52.84 | 54.58 | 55.11 | 49.25 | 49.59 | 47.86 | 52.93 |

Table 2: Average outcomes by ethnic subgroups and treatment status in 2012 and 2015

food make up a large share of household expenditures (about 50%) and a high proportion of expenditure corresponds to self-consumption (about 23%). We use the logarithm of total household expenditures net of health expenditures, equivalized by the square root of the number of members in the household to correct for differences in family composition. Between both waves, the average household size in our sample is reduced from 3.01 to 2.85 persons (with slightly smaller households for the Aymara subpopulation). As is clear from Table 2, expenditures increase for the treated Mestizo and Quechua, but not for the Aymara.

A *health* index is constructed based on the first (polychoric) principal component of a battery of five sub-dimensions of the 36-Item Short Form Survey (SF-36). We include physical functioning, physical role functioning, bodily pain, and general health as sub-dimensions.[13] The obtained principal component measure is normalized to an index between 0 and 100. Table 2 shows that the health of the treated Aymara has improved remarkably, despite the fact that all respondents have grown older between 2012 and 2015. We will call the remarkable health improvement for the treated Aymara, the "Aymara miracle" and we will return to it in Section 4.

A similar result is found for being free of limitations of *activities of daily life* (ADL), measured by the first (polychoric) principal component of four questions that deal with "crossing from one room to another", "eating (including cutting food, serving glasses, etc.)", "going to toilet (seating and standing from toilet)", and "getting in and out of bed". Each of these questions is measured on a 3-point scale and the resulting measure is again normalized to an index between 0 and 100. Here also we see that the situation of the treated Aymara has improved.

The fourth dimension, *respect*, captures the quality of the relationship with relatives and is measured by the first (polychoric) principal component of two questions: "Do you consider that your relatives treat you with respect?" and "Do you consider that your relatives respect your opinions and interests?", each measured on a 5-point scale. The resulting measure is normalized to an index between 0 and 100. While this outcome is rather stable for Mestizo and Quechua, it increases for the Aymara, both in the control and in the treatment group.

Let us finally look at a measure of *subjective well-being* (SWB), based on seven

[13] The SF-36, developed by the Rand institute, is widely used to assess health of adult respondents. The included dimensions form the physical health component of the SF-36. The emotional component of the SF-36 is not fully included in the ESBAM data set and not used in our analysis.

satisfaction questions (satisfaction with your health, yourself, your capacity to perform daily life activities, your personal relationships, the place where you live, your relationships with family, and your life as a whole). Each of these questions is measured on a 4-point scale. The first (polychoric) principal component of these seven questions is normalized to obtain an index between 0 and 100. We see that the Aymara have a higher subjective well-being in 2015 compared to 2012, whereas subjective well-being is rather stable for the other two ethnic groups. Subjective well-being can be interpreted as an overall measure of well-being, capturing the subjective evaluation of one's situation on all relevant life dimensions. Alternatively, it can also be seen as one specific life dimension with the same status as the other dimensions. We come back to these different interpretations in Section 5.

### 3.2 The impact of the Pension 65 program

Table 2 shows that some outcomes remain relatively stable over time while others are changing, and that the trends differ across the ethnic groups. A more careful analysis is needed to see whether these changes are affected by the Pension 65 program or not. Indeed, the changes in the outcomes may have several causes. First, they may be caused by developments in the Peruvian society between 2012 and 2015, such as economic growth or policy changes unrelated to the Pension 65 program. Second, the panel is ageing. Over the considered period, average age increases from 71 to 74. This ageing is likely to impact the health of our respondents and the number of respondents who are able to work. Third, some changes may be caused by the initial differences between the treated (extremely poor) and non-treated (non-extremely poor) Peruvians. Finally, some of the observed trends may indeed be an effect of the Pension 65 program.

To evaluate the effect of the Pension 65 program, we compare the trend of the outcomes for the control group and for the treatment group on either side of the eligibility threshold. To do that, we use a panel regression difference-in-differences approach with controls for some time-varying observable characteristics and individual fixed effects to control for time-invariant (un)observable

characteristics.[14] To be precise, let

$$o_{it} = \alpha_i + \beta z_{it} + \gamma D2_t + \eta \left( D2_t \times P65_i \right) + \varepsilon_{it}, \tag{1}$$

where $o_{it}$ denotes an outcome variable (expenditures, health, ADL, respect or subjective well-being), $z_{it}$ a vector of time dependent characteristics of individual $i$ (altitude of district, urban status, marital status, and work status), $\alpha_i$ an individual fixed effect, $\varepsilon_{it}$ an idiosyncratic error, and $D2$ and $P65$ dummies of the second wave and the treatment respectively.[15] We are mainly interested in the treatment parameter $\eta$. First differencing equation (1) yields

$$\Delta o_i = \beta \triangle z_i + \gamma + \eta P65_i + \triangle \varepsilon_i. \tag{2}$$

To capture potential heterogeneous treatment effects, we will interact the treatment parameter with ethnicity dummies. Moreover, we will mainly focus on the results interacting the time trend $\gamma$ in equation (1) with ethnicity dummies to allow for heterogeneous time trends in the ethnic groups within the control group.

The treatment parameter can be consistently estimated by OLS if the treatment status is uncorrelated with changes in the idiosyncratic errors, i.e., if $E(P65_i \triangle \varepsilon_i) = 0$. Given that the treatment decision on the basis of the SISFOH score may be correlated with some unobservable time varying characteristics, we cannot take it for granted that this assumption is satisfied.

To check the balancedness of treatment and control group, we perform a $t$-test of the differences in the averages of some relevant variables between the control and treatment group in 2012. The results in Table 3 show that the respondents in the control group have significantly larger average equivalized expenditures than the respondents in the treatment group ($p < 0.000$). As we can expect the SISFOH score (used to demarcate the treatment and control group) to be correlated with the expenditures, this result is not surprising. Other significant differences are found for the health index ($p < 0.05$), urban status ($p < 0.000$),

[14] Bando et al. (2016) evaluate the program by comparing the outcomes in the follow-up, while controlling (linearly) for the distance to the eligibility threshold and regional fixed effects. Note that in our specification the individual fixed effect controls for all time-invariant factors, among which is the distance to the eligibility threshold (which is unobservable, but time-invariant).

[15] As we don't observe when the respondents in the treatment group received the treatment for the first time, we cannot distinguish between treated respondents to whom the program was rolled out early (in 2013) or late (in 2015).

and for years of education ($p < 0.000$), again variables that can be expected to be correlated with the SISFOH scores.[16]

| | control | treatment | $t$-value | $p$-value |
|---|---|---|---|---|
| expenditures (in logarithm) | 5.81 | 5.69 | 5.29 | 0.000 |
| health (between 0 and 100) | 50.89 | 49.12 | 1.98 | 0.048 |
| ADL (between 0 and 100) | 80.25 | 79.37 | 0.81 | 0.418 |
| respect (between 0 and 100) | 83.97 | 84.32 | 0.35 | 0.727 |
| SWB (between 0 and 100) | 53.08 | 53.76 | 1.12 | 0.260 |
| altitude of district (in logarithm) | 7.34 | 7.39 | 0.78 | 0.435 |
| urban (in percentage) | 0.43 | 0.27 | 7.32 | 0.000 |
| married (in percentage) | 0.74 | 0.72 | 1.22 | 0.223 |
| working (in percentage) | 74.51 | 75.36 | 0.43 | 0.664 |
| age (in years) | 71.34 | 71.34 | 0.02 | 0.981 |
| male (in percentage) | 0.58 | 0.60 | 0.97 | 0.332 |

Table 3: $t$-test of balancedness of variables in 2012

The treatment parameter $\eta$ captures the causal effect of the Pension 65 program under the standard "parallel paths" assumption. This assumption requires that the trend of the outcome variables for the respondents in the treatment group and control group are not systematically different in absence of the Pension 65 program. With only one single wave of observations before the introduction of the Pension 65 program this assumption cannot be empirically tested. In fact, it seems plausible that the "parallel paths" assumption is violated, given the observed differences between both groups in 2012. This is precisely why we have allowed for heterogeneous time trends in the control groups of different ethnic groups. Still, sufficient care is needed in interpreting our estimates of $\eta$ as the causal effect of the treatment.

Tables 4 and 5 show the estimates of equation (2) for the five considered aspects of life. For each aspect, the first column provides an estimate of the overall treatment effect. The second column introduces interactions between the treatment dummy and the ethnicity dummies. In this case all non-treated respondents form the control group. The third column also introduces interactions between the time trend (i.e., the constant in our estimation in first differences) and the ethnicity dummies to allow for heterogeneous time trends in the ethnic groups within the control group. Comparing the three columns provides interesting

[16] Bando et al. (2016) find similar results when they test for the balancedness of the assignment to treatment and control groups.

| | Δ Expenditures | | | Δ Health | | | Δ ADL | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Treatment | 0.118*** | 0.109*** | 0.077** | 2.111** | 1.463 | 1.340 | 2.298* | 3.426** | 2.656 |
| | (0.027) | (0.031) | (0.035) | (1.020) | (1.141) | (1.234) | (1.372) | (1.537) | (1.671) |
| Treatment × Quechua | | 0.071 | 0.054 | | -0.802 | 0.787 | | -6.204*** | -2.035 |
| | | (0.044) | (0.059) | | (1.822) | (2.353) | | (2.368) | (3.155) |
| Treatment × Aymara | | -0.141* | 0.268** | | 12.129*** | 8.037** | | 7.255** | 1.416 |
| | | (0.082) | (0.110) | | (3.265) | (4.010) | | (3.646) | (5.100) |
| Δ Altitude | -0.125** | -0.127** | -0.129** | 0.901 | 0.882 | 0.899 | -6.301** | -6.155** | -6.141** |
| | (0.062) | (0.063) | (0.065) | (2.648) | (2.656) | (2.688) | (3.025) | (3.022) | (3.009) |
| Δ Urban status | 0.243* | 0.233 | 0.234* | -3.192 | -2.916 | -2.977 | -9.893* | -9.125* | -9.291* |
| | (0.142) | (0.143) | (0.142) | (4.912) | (4.924) | (4.946) | (5.459) | (5.508) | (5.500) |
| Δ Marital status | -0.135** | -0.132** | -0.139** | -0.017 | -0.053 | -0.020 | 2.342 | 2.113 | 2.114 |
| | (0.062) | (0.062) | (0.062) | (2.205) | (2.192) | (2.198) | (3.194) | (3.215) | (3.200) |
| Δ Work status | 0.000 | 0.000 | 0.000 | 0.062*** | 0.062*** | 0.062*** | 0.053*** | 0.053*** | 0.053*** |
| | (0.000) | (0.000) | (0.000) | (0.010) | (0.010) | (0.010) | (0.013) | (0.013) | (0.013) |
| Constant | -0.020 | -0.020 | 0.013 | -1.488** | -1.495** | -1.375* | -4.668*** | -4.674*** | -3.912*** |
| | (0.019) | (0.019) | (0.024) | (0.669) | (0.669) | (0.814) | (0.931) | (0.931) | (1.136) |
| Constant × Quechua | | | 0.016 | | | -1.587 | | | -4.165** |
| | | | (0.040) | | | (1.490) | | | (2.088) |
| Constant × Aymara | | | -0.409*** | | | 4.091* | | | 5.840 |
| | | | (0.074) | | | (2.325) | | | (3.564) |
| Observations | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 |
| $R^2$ | 0.020 | 0.023 | 0.044 | 0.023 | 0.032 | 0.035 | 0.012 | 0.018 | 0.023 |

Robust standard errors in parentheses. * < 0.10, ** < 0.05 , *** < 0.01

Table 4: Difference-in-differences estimates for different outcomes

| | Δ Respect | | | Δ SWB | | |
|---|---|---|---|---|---|---|
| | (10) | (11) | (12) | (13) | (14) | (15) |
| Treatment | 0.316 | -0.962 | -0.771 | 0.561 | -0.846 | 0.495 |
| | (1.358) | (1.451) | (1.569) | (0.693) | (0.734) | (0.809) |
| Treatment × Quechua | | 1.068 | 3.137 | | 2.119* | -0.478 |
| | | (2.405) | (3.158) | | (1.236) | (1.622) |
| Treatment × Aymara | | 14.012*** | 4.864 | | 11.917*** | 6.161** |
| | | (4.444) | (6.074) | | (2.310) | (2.885) |
| Δ Altitude | -4.211* | -4.288* | -4.239* | 0.136 | 0.038 | 0.092 |
| | (2.456) | (2.445) | (2.464) | (1.721) | (1.732) | (1.669) |
| Δ Urban status | -13.405** | -13.295** | -13.372** | -3.736 | -3.770 | -3.659 |
| | (5.398) | (5.400) | (5.409) | (3.408) | (3.417) | (3.439) |
| Δ Marital status | 2.973 | 3.004 | 3.113 | 0.171 | 0.241 | 0.404 |
| | (3.329) | (3.316) | (3.303) | (1.410) | (1.390) | (1.379) |
| Δ Work status | -0.023* | -0.023* | -0.024* | 0.022*** | 0.021*** | 0.021*** |
| | (0.013) | (0.013) | (0.013) | (0.007) | (0.006) | (0.006) |
| Constant | -1.738* | -1.745* | -1.939* | -0.486 | -0.492 | -1.825*** |
| | (0.942) | (0.942) | (1.106) | (0.469) | (0.469) | (0.578) |
| Constant × Quechua | | | -2.065 | | | 2.598** |
| | | | (2.047) | | | (1.053) |
| Constant × Aymara | | | 9.148** | | | 5.757*** |
| | | | (4.133) | | | (1.723) |
| Observations | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 |
| $R^2$ | 0.005 | 0.011 | 0.016 | 0.007 | 0.025 | 0.032 |

Robust standard errors in parentheses. * < 0.10, ** < 0.05 , *** < 0.01

Table 5: Difference-in-differences estimates for different outcomes (continued)

additional insights in the role of ethnicity on the treatment.

All estimates are obtained with four control variables that capture some time varying characteristics: $\triangle$ altitude, $\triangle$ urban status, $\triangle$ marital status, and $\triangle$ work status. Because of the regional concentration of the ethnic groups (with the Aymara in the Altiplano region of Southern Peru, see Figure 1), we introduce the logarithm of the altitude of the district in which the respondent is living as a control variable. This variable proxies for climatic and agricultural conditions, and for the quality of the public infrastructure including health care provision. Although these control variables are significant in a few cases (e.g., moving to a region at a higher altitude is negatively correlated with the change in expenditures and in ADL, and becoming "unmarried" –likely becoming a widow or widower– increases equivalized expenditures), dropping the controls from the regressions does not change the estimates of the treatment effects substantially.

Not surprisingly, being a beneficiary of the Pension 65 program does have a positive effect (of slightly more than 10%) on equivalized expenditures (see column (1) of Table 4). Yet, when looking at column (2) of Table 4, we see that there is no such treatment effect for the Aymara in comparison to the entire control group. Indeed, the treatment effect for the Aymara (-0.032) is not significantly different from 0 at any usual significance level ($p > 0.68$). However, as column (3) shows, there is a significant decrease in expenditures for the Aymara in the control group. Compared to this refined control group, the treatment effect for the treated Amayra becomes 0.344, which is significant at any usual significance level ($p < 0.002$). Whereas we witness lower expenditures in 2015 compared to 2012 for the Aymara in the control group, the Aymara in the treatment group were able to keep their expenditures more or less at the same level.[17]

More striking is the treatment effect on health: it is positive, but the simple model without interactions of column (4) masks an interesting heterogeneous effect across ethnic groups. In fact, we see in column (5) that the effect of the treatment on health is large and significant for the Aymara ($p < 0.000$) compared to the total control group, but not significantly positive for the other ethnic groups.[18] In column (6) we notice that also the health of the Aymara in the control group has increased, albeit at a much lower rate than for the Aymara

[17] Based on the ESBAM data, it is impossible for us to distinguish whether consumed quantities have dropped, or whether prices have fallen in the districts of the Altiplano region of Southern Peru where the Aymara are concentrated.

[18] Also Van de gaer et al. (2013) find strong health effects of the Mexican cash transfer program on the health opportunities of children with an indigenous background.

in the treatment group.

Also for ADL (see columns (7) to (9) of Table 4), the positive overall effect of the intervention masks some heterogeneity. The effect on ADL is positive for Mestizo, close to zero for the Quechua ($p > 0.19$) and positive for the Aymara ($p < 0.003$) in comparison to the total control group. Only for the Aymara is the treatment effect large enough to compensate for the negative effect of ageing on ADL, as measured by the estimate of the constant $\gamma$ in equation (2). When we introduce heterogeneous trends for the control groups (column (9)), most treatment effects are not precisely estimated.

In column (10) to (12) of Table 5, we see a pattern for the respect variable that is similar to the one of ADL. Column (11) shows a large effect of the intervention for Aymara that is significantly different from zero ($p < 0.01$) when comparing to the total control group, and an insignificant effect for the other groups. Yet, as we have seen already in Table 2, respect also increases for the Aymara control group, in contrast to the Quechua and the Mestizo. This overall increase of respect for all Aymara could be explained by the position taken by elderly in the Aymara culture (remember that our sample grows three years older between 2012 and 2015). Moreover, the increased "respect from relatives" for the Aymara can also be explained by the fact that the treated groups give more material and economic support to persons close to them in their personal network. This hypothesis is supported by the results of Canavire Bacarreza et al. (2017) who find that indigenous grandparents in Bolivia spend a larger part of the unconditional transfer that they receive into educational expenditures for their grandchildren.

As mentioned before, subjective well-being can be seen as an additional life dimension or as a global measure of individual well-being, integrating the other dimensions. Our results for life satisfaction (column (13) to (15) of Table 5) suggest that the Pension 65 program does not have a significantly positive effect on the Mestizo, a (small) insignificant effect on the Quechua ($p > 0.30$), and a large significant effect on the Aymara ($p < 0.000$) in comparison to the total control group. However, column (15) shows that there are heterogeneous trends in the different ethnic groups in the control group as well. Only for the Aymara the treatment effect is positive and significant ($p < 0.02$) when taking this heterogeneity into account.

Until now we looked at the different dimensions in isolation. In reality they are of course interrelated. Tables 6 and 7 show the results taking these in-

| | (1) Δ Expend. | (2) Δ Expend. | (3) Δ Expend. | (4) Δ Health | (5) Δ Health | (6) Δ Health | (7) Δ ADL | (8) Δ ADL | (9) Δ ADL |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | 0.117*** | 0.113*** | 0.076** | 1.040 | 0.530 | 0.233 | 1.072 | 2.934** | 1.905 |
| | (0.027) | (0.031) | (0.034) | (0.852) | (0.966) | (1.051) | (1.182) | (1.344) | (1.466) |
| Treatment × Quechua | | 0.060 | 0.055 | | 0.401 | 1.494 | | -6.035*** | -2.541 |
| | | (0.044) | (0.059) | | (1.484) | (1.938) | | (1.986) | (2.663) |
| Treatment × Aymara | | -0.162** | 0.240** | | 5.630** | 5.094* | | -3.851 | -4.671 |
| | | (0.081) | (0.110) | | (2.488) | (3.088) | | (2.988) | (4.182) |
| Δ Expenditures | | | | 1.037 | 1.110 | 1.158 | -1.976** | -1.911** | -1.797* |
| | | | | (0.706) | (0.700) | (0.707) | (0.936) | (0.936) | (0.950) |
| Δ Health | 0.001 | 0.001 | 0.001 | | | | 0.634*** | 0.632*** | 0.629*** |
| | (0.001) | (0.001) | (0.001) | | | | (0.027) | (0.028) | (0.028) |
| Δ ADL | -0.001** | -0.001** | -0.001* | 0.325*** | 0.325*** | 0.324*** | | | |
| | (0.001) | (0.001) | (0.001) | (0.014) | (0.014) | (0.014) | | | |
| Δ Respect | -0.001** | -0.001* | -0.001 | 0.019 | 0.017 | 0.017 | 0.065*** | 0.066*** | 0.064*** |
| | (0.000) | (0.000) | (0.000) | (0.014) | (0.014) | (0.014) | (0.020) | (0.020) | (0.021) |
| Δ SWB | 0.002* | 0.002* | 0.002** | 0.348*** | 0.340*** | 0.341*** | 0.180*** | 0.188*** | 0.192*** |
| | (0.001) | (0.001) | (0.001) | (0.030) | (0.030) | (0.031) | (0.041) | (0.042) | (0.042) |
| Δ Altitude | -0.131** | -0.132** | -0.133** | 3.108 | 3.080 | 3.074 | -6.858*** | -6.666*** | -6.674*** |
| | (0.054) | (0.055) | (0.057) | (1.968) | (1.973) | (1.992) | (2.216) | (2.227) | (2.170) |
| Δ Urban status | 0.109 | 0.102 | 0.106 | 1.460 | 1.442 | 1.379 | -6.089 | -5.482 | -5.655 |
| | (0.117) | (0.117) | (0.115) | (4.055) | (4.052) | (4.076) | (4.608) | (4.627) | (4.600) |
| Δ Marital status | -0.128** | -0.126** | -0.134** | -0.758 | -0.730 | -0.737 | 1.867 | 1.654 | 1.603 |
| | (0.061) | (0.061) | (0.060) | (1.661) | (1.652) | (1.655) | (2.675) | (2.671) | (2.668) |
| Δ Work status | 0.000 | 0.000 | 0.000 | 0.037*** | 0.037*** | 0.037*** | 0.012 | 0.013 | 0.012 |
| | (0.000) | (0.000) | (0.000) | (0.008) | (0.008) | (0.008) | (0.011) | (0.011) | (0.011) |
| Constant | -0.024 | -0.024 | 0.013 | 0.254 | 0.246 | 0.535 | -3.565*** | -3.560*** | -2.549** |
| | (0.019) | (0.019) | (0.024) | (0.566) | (0.566) | (0.699) | (0.795) | (0.795) | (0.992) |
| Constant × Quechua | | | 0.004 | | | -1.103 | | | -3.510** |
| | | | (0.040) | | | (1.256) | | | (1.775) |
| Constant × Aymara | | | -0.410*** | | | 0.546 | | | 0.848 |
| | | | (0.075) | | | (1.857) | | | (3.034) |
| Observations | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 |
| $R^2$ | 0.023 | 0.027 | 0.047 | 0.330 | 0.331 | 0.332 | 0.280 | 0.284 | 0.286 |

Robust standard errors in parentheses. * < 0.10, ** < 0.05 , *** < 0.01

Table 6: Difference-in-differences for outcomes including changes in other outcomes as controls.

| | (1) Δ Respect | (2) Δ Respect | (3) Δ Respect | (4) Δ SWB | (5) Δ SWB | (6) Δ SWB | (7) Δ SWB | (8) Δ SWB | (9) Δ SWB |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | 0.156 | -0.890 | -1.020 | -0.118 | -1.386** | 0.032 | -0.202 | -1.343** | 0.045 |
| | (1.341) | (1.449) | (1.568) | (0.639) | (0.684) | (0.757) | (0.636) | (0.684) | (0.757) |
| Treatment × Quechua | | 1.302 | 3.476 | | 2.492** | -0.741 | | 2.480** | -0.721 |
| | | (2.325) | (3.090) | | (1.137) | (1.501) | | (1.138) | (1.501) |
| Treatment × Aymara | | 9.858** | 3.494 | | 8.586*** | 3.964 | | 7.124*** | 2.770 |
| | | (4.391) | (5.862) | | (1.989) | (2.449) | | (1.917) | (2.371) |
| Δ Expenditures | -2.287** | -2.158* | -1.776 | 0.975* | 1.044* | 1.213** | 1.010* | 1.051* | 1.210** |
| | (1.128) | (1.130) | (1.129) | (0.551) | (0.547) | (0.554) | (0.552) | (0.549) | (0.554) |
| Δ Health | 0.049 | 0.044 | 0.042 | 0.201*** | 0.195*** | 0.194*** | 0.184*** | 0.183*** | 0.182*** |
| | (0.036) | (0.036) | (0.036) | (0.018) | (0.017) | (0.017) | (0.018) | (0.018) | (0.018) |
| Δ Health × Aymara | | | | | | | 0.190*** | 0.147*** | 0.140*** |
| | | | | | | | (0.049) | (0.047) | (0.047) |
| Δ ADL | 0.084*** | 0.085*** | 0.083*** | 0.053*** | 0.055*** | 0.056*** | 0.052*** | 0.054*** | 0.055*** |
| | (0.027) | (0.027) | (0.027) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) |
| Δ Respect | | | | 0.055*** | 0.051*** | 0.051*** | 0.051*** | 0.049*** | 0.048*** |
| | | | | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) |
| Δ SWB | 0.239*** | 0.226*** | 0.223*** | | | | | | |
| | (0.051) | (0.051) | (0.051) | | | | | | |
| Δ Altitude | -4.031 | -4.069 | -4.007 | 0.638 | 0.552 | 0.626 | 0.501 | 0.451 | 0.528 |
| | (2.594) | (2.569) | (2.584) | (1.605) | (1.615) | (1.551) | (1.648) | (1.649) | (1.581) |
| Δ Urban status | -11.248** | -11.293** | -11.460** | -1.950 | -2.132 | -2.022 | -1.996 | -2.177 | -2.063 |
| | (5.426) | (5.412) | (5.417) | (3.104) | (3.108) | (3.142) | (3.110) | (3.115) | (3.145) |
| Δ Marital status | 2.433 | 2.490 | 2.605 | 0.016 | 0.116 | 0.298 | -0.083 | 0.036 | 0.216 |
| | (3.277) | (3.272) | (3.267) | (1.173) | (1.170) | (1.161) | (1.176) | (1.176) | (1.166) |
| Δ Work status | -0.035*** | -0.035*** | -0.035*** | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 |
| | (0.013) | (0.013) | (0.013) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) |
| Constant | -1.205 | -1.213 | -1.127 | 0.177 | 0.169 | -1.257** | 0.080 | 0.097 | -1.297** |
| | (0.931) | (0.931) | (1.108) | (0.441) | (0.441) | (0.548) | (0.440) | (0.440) | (0.548) |
| Constant × Quechua | | | -2.210 | | | 3.228*** | | | 3.197*** |
| | | | (2.053) | | | (0.993) | | | (0.993) |
| Constant × Aymara | | | 6.490 | | | 4.664*** | | | 4.461*** |
| | | | (3.948) | | | (1.458) | | | (1.430) |
| Observations | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 | 1968 |
| $R^2$ | 0.042 | 0.045 | 0.048 | 0.159 | 0.169 | 0.176 | 0.165 | 0.172 | 0.179 |

Robust standard errors in parentheses. * < 0.10, ** < 0.05 , *** < 0.01

Table 7: Difference-in-differences for outcomes including changes in other outcomes as controls.

terrelationships into account. Expenditures are introduced after a logarithmic transformation. For expenditures, health, ADL and respect, most results are as expected. Examples are the positive effects of ADL on health and respect, of respect on ADL, and of health and ADL on each other. More surprising is the negative association between ADL and expenditures. Overall, the results for the "reduced" treatment effects in tables 4-5 are confirmed.

Most interesting are the results for SWB. An increase in SWB is associated with an improvement in health, ADL and respect. Of course, all four dimensions (expenditures, health, ADL and respect) have a positive effect on SWB. We will show in section 4.2 how we can derive useful information about the ordinal preferences of the individuals from this kind of regression. From a normative point of view it is important to check whether the Pension 65 program has an independent effect on SWB, after controlling for the other life dimensions. Column (6) shows that this is not the case: for the Quechua as well as for the Aymara, SWB increases in the control group. This may be explained by the age effect, that is well known in the happiness literature, and it may also reflect the position of the elderly in the indigenous groups. However, the Pension 65 treatment has no significant effect on SWB after controlling for the other dimensions. In the evaluation of the program, the SWB outcomes can be interpreted as a global measure of individual well-being, but they are not relevant as a separate dimension.

All-in-all, our analysis confirms that introducing ethnic differences reveals interesting heterogeneous effects of the program that otherwise remain hidden.[19] In Section 5 we will return to the normative implications of these findings.

# 4 Is there an Aymara miracle?

The findings in the previous section show that Pension 65 seems to have a positive effect on the health of the Aymara. For the other ethnic groups we witness increased expenditures and hardly any effect on the health index. Is

[19] Bando et al. (2016) also find that the Pension 65 program has a positive effect on total expenditures. They do not subtract health expenses to get at a net expenditures concept. They find no effect on (an arguably ad-hoc measure of) physical health. Our health index is defined more broadly and we take ethnic differences into account in a restricted sample. Since the positive effect on health in our result is driven by the performance of the Aymara, it is not surprising that a positive health effect is no longer found in their larger sample (where the Aymara are an even smaller minority). Contrary to our results, Bando et al. (2016) also find a positive effect on subjective well-being. However, they measure subjective well-being by a depression scale, which is closer to health than our life satisfaction measure.

it possible to explain these differences? We will first check the robustness of the remarkable increase in health for the Aymara. Then we will investigate whether the differences may be explained by preference heterogeneity between the Aymara and the other ethnic groups, i.e., by a different view on the meaning of a good life.

## 4.1 Robustness of the Aymara miracle

In this section we present several robustness checks of the health increase of the Aymara subpopulation. First, one may wonder whether this finding is a mere artefact of the way in which we have constructed the health index. Recall that we have used (polychoric) principal component weights in the health index. Yet, as Table 14 in Appendix shows, the Aymara score better on each of the four considered sub-dimensions of the SF-36, so the result is not sensitive to the weights given to these sub-dimensions in the constructed health index.

Moreover, a more detailed analysis of the ESBAM data set shows that the same sharp increase in health outcomes is found for other indicators as well. The last column of Table 14 shows that the Aymara suffer less from nutritional deficiency as measured by the "mini nutritional assessment" (MNA), whereas the score does not significantly change for the Quechua ($p > 0.80$).

Note, furthermore, that Table 4 (column (6)) shows that the Aymara miracle extends to the ADL-index, which captures physical fitness. The consistency between the results with the health and ADL indices softens -to some extent- our worries about the lack of balance with respect to health in the assignment between control and treatment groups (as documented in Table 3). Indeed, this lack of balance is not found for the ADL-index.

Third, given the small sample of Aymara respondents, one may be concerned that our results are driven by a few outliers. Yet, similar patterns are found for the median and all other percentile values. Figure 2 performs a test of first order stochastic dominance on health improvement (health index in 2015 minus health index in 2012) between the six groups identified in Table 1. The treated Aymara dominate all other groups when it comes to health improvement, and in particular the Aymara in the control group. We can therefore reject the hypothesis that the sharp increase in health outcomes of the Aymara is driven by a few outliers.

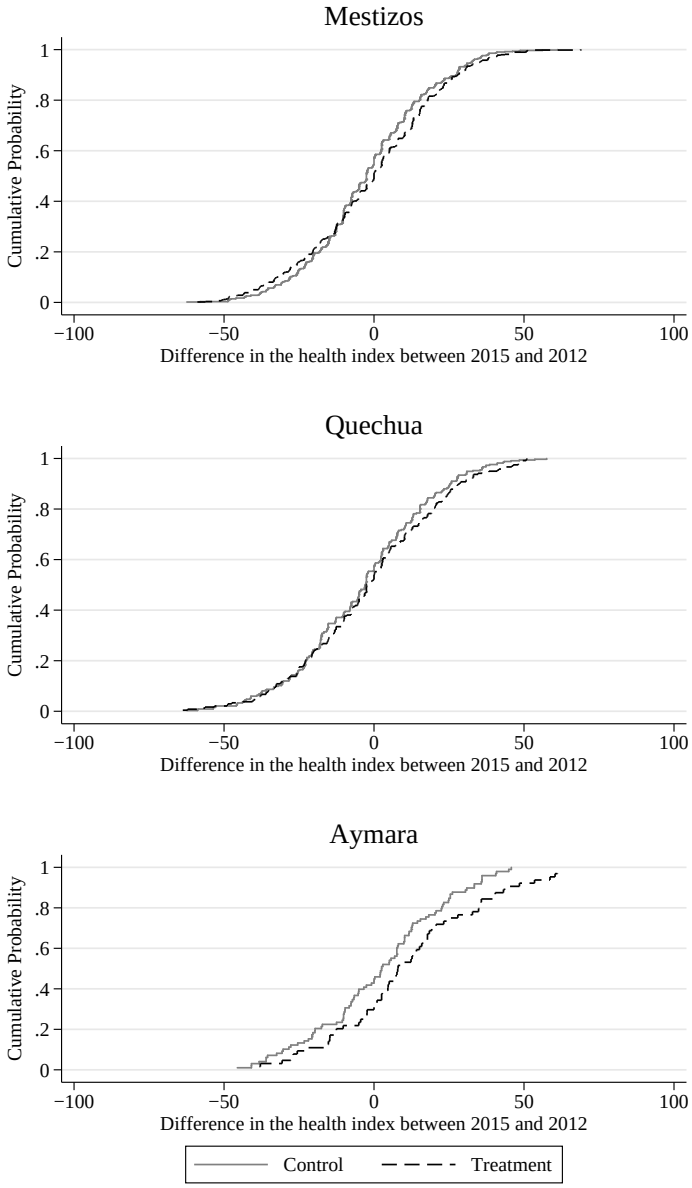


Figure 2: Cumulative Distribution Function for the difference in the health index between 2015 and 2012

## 4.2 A different view of the good life?

The standard economic approach to health differences focuses on opportunities and preferences. This framework offers a natural explanation to the Aymara miracle. Health cannot be bought directly on a market, but has to be produced by investments in health care, in time allocated to physical activities, in nutrition, etc. These investments have an opportunity cost in terms of expenditures. In the popular Grossman model this leads to a downward sloping health production frontier that is represented by the full lines in Figure 3 (see, e.g., the analysis in Zweifel et al. (2009)).

For the sake of illustration we assume that the health production possibilities for the Aymara in 2012 (depicted by the full black line closest to the origin in Figure 3), who are living in harsher circumstances in the Altiplano region of Southern Peru, are more restricted than for the Mestizo in 2012 (the full grey line closest to the origin).[20] Let us now assume that the Aymara and the Mestizo have a different view on the good life: the Aymara have steeper indifference curves, meaning that they find health relatively more important. This leads to a 2012 situation where the health of the Aymara and the Mestizo is similar, but the Mestizo have larger expenditures. The introduction of the Pension 65 program shifts both health production frontiers outwards. In Figure 3, this leads to an increase in consumption and a (slight) decrease in health for the Mestizo, and to a large increase in health accompanied by a slight decrease in expenditures for the Aymara. This schematic representation mimics what we have found in our data (see Table 2, for instance). In this interpretation, the interaction between differences in preferences and differences in the production frontier is the driving force behind our findings. Preference differences play an essential role.

We do not have the data to estimate a full structural model of health production by the households. We can, however, derive direct information about preferences from the estimation of a life satisfaction regression. At least since Clark and Oswald (2002) and Van Praag and Baarsma (2005), it has become standard practice to derive estimates of willingness-to-pay for non-market goods, i.e., marginal rates of substitution, from satisfaction equations.[21] Decancq et al. (2015a) discuss the crucial consistency assumption that is needed for this. This

[20] Evidence shows that average life expectancy is 30 years shorter in the highlands than in Lima, for instance (World Bank, 2015).

[21] See also Decancq et al. (2015a); Decancq and Schokkaert (2015); Decancq and Neumann (2016); Decancq and Michiels (2018) for applications to the measurement of well-being. Fujiwara and Dolan (2016) provide a critical discussion.

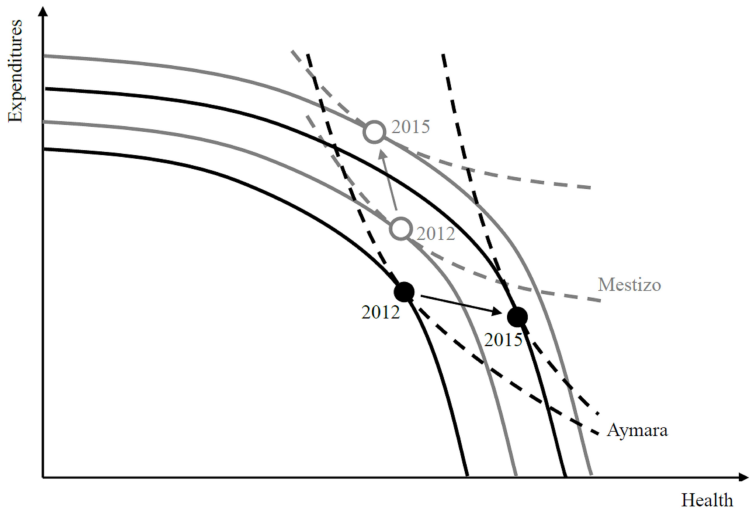


Figure 3: Preference differences and the Aymara miracle

consistency assumption requires that the subjective well-being index tracks the preferences of the respondents. Each respondent is assumed to report a higher subjective well-being index in situations that she prefers (which reflect a better life according to her). This consistency question seems plausible, but is hard to test empirically.

For the sake of convenience, we rewrite eq. (1) for subjective well-being after having introduced the other life dimensions:

$$SWB_{it} = \alpha_i + \beta ln(y_{it}) + \gamma x_{it} + \delta z_{it} + \zeta D2_t + \eta(D2_t \times P65_i) + \varepsilon_{it}, \qquad (3)$$

where $SWB_{it}$ is the subjective well-being index of respondent $i$ in period $t$, $y_{it}$ is the expenditures variable and $x_{it}$ are the non-income dimensions (health, ADL, and respect). The vector $z_{it}$ contains some time-dependent control variables (altitude of district, urban status, marital status, and work status), $D2_t$ is a dummy indicating the follow-up wave, $P65_i$ is the treatment dummy and $\varepsilon_{it}$ an idiosyncratic error. We are interested in retrieving information about the preferences over expenditures $y_{it}$ and non-income dimensions $x_{it}$, i.e., in the coefficients $\beta$ and $\gamma$ which determine the marginal rates of substitution. The other terms in equation (3) are included to capture the interpersonal differences in the use of the response scale.[22] In particular, the individual fixed effects $\alpha_i$

[22] In a context similar to ours, Beegle et al. (2012) and Ravallion et al. (2016) use a vignette

play a crucial role to control for time-invariant personality traits, as discussed by Ferrer-i-Carbonell and Frijters (2004). First differencing equation (3) yields

$$\Delta SWB_i = \beta \triangle ln(y_i) + \gamma \triangle x_i + \delta \triangle z_i + \zeta + \eta P65_i + \triangle \varepsilon_i. \qquad (4)$$

This is an extended version of eq. (2) that, as explained in section 3.2, can be estimated with standard OLS. Ordinal preference differences between groups, to be distinguished from differences in the response scale, can be captured by including interactions between the dimensions of well-being (expenditures, health, ADL, and respect) and the three ethnic groups. To arrive at a parsimonious model, we have started with a full set of all possible interactions, and then we dropped the least significant interaction term from the model. We repeat this procedure until all remaining interactions are significant at the 10% level. The results were already shown in columns (7)-(9) of Table 7. As can be seen from column (9), only one interaction remains at the end: that is the interaction between the health index and the Aymara dummy.

The results in Table 7 are in line with results found in other satisfaction studies, including those with Peruvian data.[23] We focus on the results with differentiated trends for the different control groups. All dimensions of life have a significantly positive effect on the subjective well-being index. The significant interaction effect for the Aymara suggests that they attach more importance to health than the other ethnic groups in our sample.[24] This difference in preferences is further illustrated by the results in Table 8, which shows the willingness-to-pay for a one point increase on a non-income dimension (measured on a 100-point scale), as a percentage of expenditures.[25] Respondents who identify as Mestizo and Quechua stay on the same indifference curve when their health increases by one point and their expenditures reduce by 15.06 percent. Respondents who identify as Aymara, on the other hand, are willing to reduce their expenditures much

study and show that, although subjective life satisfaction answers are sensitive to interpersonal differences in scale use, the estimated trade-offs between various life dimensions are reassuringly robust.

[23] GuillenRoyo (2008; 2011) emphasizes the importance of reference group consumption for life satisfaction. Our data are not rich enough to estimate such reference group effects. Interestingly, however, she finds that reference groups do not matter for appraisal in the domain of physical health, which is largely made on the basis of the objective situation of the household.

[24] Using alternative weighting schemes to the (polychoric) principal component weights used to construct the well-being index lead to similar results, even when satisfaction with health is removed from the set of considered domain satisfactions.

[25] From equation (3) it follows that the willingness-to-pay can be computed as $100 \times \gamma/\beta$.

more (26.60 percent) to stay on the same indifference curve after an increase of their health index. Clearly, the marginal rate of substitution between health and expenditures is much larger for the Aymara than for the other groups. The resulting steeper indifference curves are consistent with the hypothesized explanation in Figure 3.

| | Mestizo | Quechua | Aymara |
|---|---|---|---|
| Health | 15.06 | 15.06 | 26.60 |
| ADL | 4.58 | 4.58 | 4.58 |
| Respect | 3.96 | 3.96 | 3.96 |

Table 8: WTP for an increase of 1 point on the index (expressed as percentage of expenditures)

The fact that the Aymara seem to have different preferences for health can be interpreted in many ways. One possibility is that these preferences reflect deeper underlying cultural differences, with the ideal of *suma jakaña* embodying more attention for the non-material dimensions of life (Calestani, 2009; Artaraz and Calestani, 2015). Another interpretation of the Aymara preference for health could be that health (and a higher level of physical fitness) has a larger instrumental value if one is living in harsher climatic circumstances and in an environment with a less developed infrastructure (as is the Altiplano region of Southern Peru).[26] These two interpretations can go perfectly hand-in-hand though.

Apart from the direct effect of the improved access to health care, there are other channels through which a larger preference for health may affect the impact of the Pension 65 program. The income transfer may allow beneficiaries a shift to a healthier life style. As mentioned before, Table 14 in the Appendix shows that we see among the treated Aymara lower levels of nutritional deficiency, as measured by their MNA score, which could reflect a shift to food with a higher nutritional quality. Moreover, the improved access to health care may induce a greater awareness of health problems and an improvement of the information about health.[27] This latter channel is proposed by Bernal et al. (2017) as the

[26] Already in the 19th century, Forbes (1870, p. 224) discusses a similar mechanism and observes that Aymara in the highlands of Bolivia and Peru enjoy a particularly robust health. He argues that "One great reason for this, however, is that, owing to the great mortality which takes place amongst the infants, a sort of natural selection asserts itself, and only the very strong children survive the first few years after birth."

[27] In line with the findings of Bando et al. (2016, p. 11), however, we find no significant

main explanation for the impact of expanding health insurance coverage through the Peruvian "Integral Health Insurance Plan".

# 5 Ethnicity, preference differences and program evaluation

The standard approach to evaluate a program is to look at the outcomes in one or several dimensions of well-being in isolation. Typically, the focus is on a monetary variable such as expenditures or income, but it can also be a health outcome or any other unidimensional outcome measure, see Bando et al. (2016) and Galiani et al. (2016) for examples. We have seen in Table 4 that the expenditures of the Mestizo subpopulation have increased after the Pension 65 program, whereas the expenditures of the Aymara have not. For health the opposite is true: the health index of the individuals who identify as Aymara has increased, whereas the health index of the other individuals has remained stable. Clearly, whether we focus on expenditures or on health matters for the evaluation of the program for the different subgroups.

It is not obvious how to evaluate the overall effect of a program such as Pension 65 when it has different effects on the overall well-being of different subgroups, *a fortiori* when these subgroups have different preferences. In this section, we will discuss two normative criteria for program evaluation that are sensitive to preferences. These criteria are inspired by the recent literature on well-being measurement.[28]

## 5.1 Two normative criteria for program evaluation

A first, obvious, possibility is to focus on a measure of subjective well-being. This leads to the following normative criterion for program evaluation:

**Subjective well-being criterion** *A program benefits a recipient if she reports a higher subjective well-being after the introduction of the program.*

This criterion has appealing features. It is sensitive to preference differences between ethnic groups or individuals in general. Indeed, if there are ethnic differences in the valuation of the different dimensions of life, they will be reflected

effect of the Pension 65 program on the use of health services.

[28] A elaborate discussion of different approaches to measuring well-being can be found in Fleurbaey and Blanchet (2013); Decancq et al. (2015b); Adler and Fleurbaey (2016).

in the measure of subjective well-being (under the consistency assumption that the measure of subjective well-being tracks the preferences of the respondents). Table 7 gives an example of how a measure of subjective well-being captures the valuation of the different dimensions of life. Moreover, the criterion is easy to implement. In the notation of the previous section, testing this criterion boils down to checking whether the subjective well-being index has increased, i.e., to check whether $\Delta SWB_i > 0$.

However, the subjective well-being criterion has also some normative drawbacks. The most important is that measures of subjective well-being do not only capture the outcomes in the different dimensions of life as weighted by preferences, but they also reflect the differences that are due to changing aspirations and expectations, or to mere individual idiosyncrasies in the use of the response scale.[29] Figure 3 illustrates. Suppose a person who is at the black point denoted "2012" before the introduction of a program and in the black point "2015" after the introduction of the program. She has moved to a higher indifference curve and prefers the situation after the introduction to the situation before the introduction of the program. However, it is possible that her subjective well-being is lower in the black point "2015" compared to "2012". She can be confronted with other people that do even better than her, for instance, or maybe something miserable has happened in her personal life. These events may have shifted her frame of reference so that she reports a lower subjective well-being, despite the fact that she moved to a higher indifference curve. Indeed, both in 2012 and in period 2015 (when she had larger aspirations) she prefers the situation after the introduction of the program to the initial situation. A similar story can be told when the program would bring the beneficiary from the black point "2015" to the point "2012": expenditures remain the same, but health deteriorates. Since the individual prefers a better health, the program has brought her to a worse bundle according to her preferences. Yet, it is possible that she adapts to her new situation with lower health and reaches the same subjective well-being level as before: she would still prefer to return to the situation "2015", but her subjective well-being has returned to its original level.

Ultimately, it is a normative question whether we want to take up differences in aspirations when evaluating a program. If one thinks that the objective

[29] Our argument echoes the discussion of "physical-condition neglect" by Sen (1985, p. 21). Sen argues "A person who is ill-fed, undernourished, unsheltered and ill can still be high up in the scale of happiness or desire-fulfillment if he or she has learned to have 'realistic' desires and to take pleasure in small mercies".

situation as evaluated by the ordinal preferences rather than the subjective evaluation is what matters to evaluate a program, the subjective well-being criterion is not attractive. The following preference-based criterion presents a natural alternative.

**Preference-based criterion** *A program benefits a recipient if she is lifted to a higher indifference curve after the introduction of the program.*

This criterion is indeed very natural, to the extent that it will seem almost tautological for economists. It is ethically very robust: all well-being measures that are based on the information of ordinal preferences alone, are consistent with the proposed preference-based criterion. Whenever the preference-based criterion classifies a program as beneficial, the so-called equivalent income measure will increase, for instance, as well as all quantity-metric and money-metric well-being measures as discussed by Deaton and Muellbauer (1980, pp. 179-182)).[30] What all these well-being measures essentially do is attach interpersonally comparable numerical labels to the indifference curves of the respondents, so that higher well-being is observed on higher indifference curves. As there is no consensus in the literature about the attractiveness of different ways to attach labels to indifference curves, we take a more robust approach here and limit ourselves to ordinal preference information. This robustness comes at a price, however. The proposed criterion is able to tell us whether the program was beneficial or not, but it does not provide us with an exact quantification of how much the program has benefited the person at hand. For the purpose of our analysis here, such a quantification is not needed.

The analysis in the previous section suggests an easy test for the criterion.[31] Individual $i$ will reach a higher indifference curve if $(\beta \triangle ln(y_i) + \gamma \triangle x_i) > 0$. Equation (3) illustrates the difference between the subjective well-being criterion and the preference-based criterion. The subjective well-being criterion looks at the evolution of the subjective well-being index over time. This evolution is sensitive to changes in observable controls, the time trend, and changes in the idiosyncratic error term, as well as the ordinal preferences. The preference-based criterion, on the other hand, is only sensitive to the ordinal preferences of the respondents.

[30] On equivalent incomes, see Decancq et al. (2015a, 2017). Samuelson (1977); Deaton (1979); Fleurbaey and Tadenuma (2014) provide examples of quantity metric well-being measures, Samuelson (1974); Bosmans et al. (2018) provide examples of money metrics. For recent surveys, see Fleurbaey and Maniquet (2011); Fleurbaey and Blanchet (2013).

[31] Alternative methods to estimate preferences are based on observed behavior (revealed preferences) or contingent valuation or discrete choice experiments (stated preferences).

We mentioned earlier the interesting possibility to include individual happiness as a separate life dimension. Sen (1985, p. 17) argues that "It would be odd to claim that a person broken down by pain and misery is doing very well". However, with the approach to estimate preferences used here, determining the relative weight of happiness (the marginal rate of substitution between happiness and the other dimensions of life) would only be possible if we had an independent measure of "feelings of happiness" that does not coincide (or is not spuriously correlated) with the subjective well-being index. It seems unlikely that such a measure of emotions could be constructed on the basis of a direct survey question alone. Alternatives could be based on experience sampling or day reconstruction (Kahneman and Krueger, 2006) or preferences over feelings of happiness could be directly elicited in binary choices, see Benjamin et al. (2012); Adler et al. (2017). Unfortunately, neither a direct measure of feelings of happiness, nor a battery of binary choices is available in the ESBAM data. Moreover, as we have seen, the results in Table 7 suggest that the treatment has no independent effect on life satisfaction, after controlling for the other dimensions of life.

## 5.2 Evaluating the Pension 65 program

We now apply the normative criteria to the Pension 65 program. Using the coefficient estimates from column (9) of Table 7, Table 9 shows the share of the sample for which the program was beneficial, broken down by ethnic group and treatment status. The first four rows show simply the share of the sample whose outcomes have increased. The results are in line with the earlier findings of this paper. A majority of the Mestizo and Quechua sub-populations have witnessed an increase in their expenditures, which is not the case for the Aymara. The opposite is true for the health index, where a large majority of Aymara saw an increase in their health, contrary to the other groups (except for the treated Mestizo). A majority of the Mestizo and Aymara subgroup saw an increase of their ADL index, while there is a majority of Quechua whose ADL index worsened. The respect index increased for a (large) majority in all subgroups, presumably due to the ageing of the sample and the important role played by elderly in the Peruvian society.

The next two rows of Table 9 focus on the summary criteria. Interestingly, we see that a majority of all groups report a higher subjective well-being, whereas only for the Amayra a majority reaches a higher indifference curve. When we focus on the control group of the Quechua, we see a remarkable difference between both

| Criterion | Total | Mestizo | | Quechua | | Aymara | |
|---|---|---|---|---|---|---|---|
| | | C | T | C | T | C | T |
| Expenditures | 53.76 | 51.97 | 57.86 | 50.30 | 64.02 | 32.65 | 45.31 |
| Health | 48.73 | 45.51 | 51.17 | 44.31 | 48.12 | 57.14 | 70.31 |
| ADL | 55.69 | 57.32 | 61.71 | 46.41 | 43.93 | 61.22 | 67.19 |
| Respect | 68.60 | 70.55 | 71.57 | 60.48 | 63.18 | 74.49 | 75.00 |
| Subjective well-being | 53.46 | 50.87 | 51.67 | 54.49 | 53.14 | 64.29 | 75.00 |
| Preference-based | 45.38 | 42.52 | 48.16 | 39.82 | 45.19 | 52.04 | 67.19 |

Table 9: Percentage of respondents who are better of in 2015 compared to 2012

measures: less than 40% of this group has moved to a higher indifference curve, but about 55% report a higher subjective well-being. This finding suggests that between 2012 and 2015 this group has substantially adjusted their frame of reference. This is consistent with the so-called "satisfaction paradox" that is found in ageing studies, see, e.g., Stone et al. (2010); Gana et al. (2012).

Table 10 zooms in on the difference between the subjective well-being criterion and the preference-based criterion. We see that 22% of the respondents report a higher level of subjective well-being, while they end up on a lower indifference curve and that 14% reach a higher indifference curve and report a lower level of life satisfaction. The main explanation of this result is the large variability in the life satisfaction answers, which may be due to changes in the frame of reference or just to individual idiosyncrasies.

| | Lower SWB | Higher SWB | Total |
|---|---|---|---|
| Lower indifference curve | 32.27 | 22.36 | 54.62 |
| Higher indifference curve | 14.28 | 31.10 | 45.38 |
| Total | 46.54 | 53.46 | |

Table 10: Percentage of respondents who moved to a higher indifference curve, compared to SWB

In order to quantify the impact of the Pension 65 program on the probability of moving to a higher indifference curve, we use a difference-in-differences method in the same spirit as the one used in Section 3. We now estimate a logit model in which the binary dependent variable takes the value of 1 when the individuals have moved to a higher or equal indifference curve, and 0 otherwise. The results are given in Table 11. We see in column (1) that, overall, the Pension 65

| | Higher indifference curve | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Treatment | 0.268*** | 0.248** | 0.250** |
| | (0.092) | (0.104) | (0.116) |
| Treatment × Aymara | | 0.785*** | 0.405 |
| | | (0.276) | (0.352) |
| Treatment × Quechua | | -0.128 | -0.022 |
| | | (0.155) | (0.208) |
| Δ Altitude | -0.297 | -0.296 | -0.296 |
| | (0.217) | (0.218) | (0.224) |
| Δ Urban status | -0.615 | -0.591 | -0.598 |
| | (0.465) | (0.469) | (0.473) |
| Δ Marital status | -0.202 | -0.207 | -0.204 |
| | (0.192) | (0.192) | (0.192) |
| Δ Work status | 0.005*** | 0.005*** | 0.005*** |
| | (0.001) | (0.001) | (0.001) |
| Constant | -0.258*** | -0.259*** | -0.261*** |
| | (0.064) | (0.064) | (0.081) |
| Constant × Quechua | | | -0.106 |
| | | | (0.140) |
| Constant × Aymara | | | 0.381* |
| | | | (0.219) |
| Observations | 1968 | 1968 | 1968 |
| Pseudo $R^2$ | 0.015 | 0.018 | 0.020 |

Robust standard errors in parentheses. * $< 0.10$, ** $< 0.05$ , *** $< 0.01$

Table 11: Difference-in-difference estimation of the effect of the Pension 65 program on the likeliness to move to a higher indifference curve (logit model)

program treatment has a significantly positive effect on its beneficiaries. Column (2) shows that the positive effect is more pronounced for the Aymara when comparing to the total control group. When allowing for heterogeneous time trends across the ethnic groups in the control group (in column (3)), we see that the Aymara in the control group have also moved to a higher indifference curve. The point estimate of the additional effect of the treatment on the Aymara is positive, but not significantly so ($p > 0.25$). Again, leaving out the large variability in the subjective well-being index that is due to changes in the frame of reference or individual idiosyncrasies does change the results.

# 6 Conclusion

We have shown that explicitly integrating the ethnic dimension in program evaluation is possible and may yield relevant new insights. The non-contributory pension program Pension 65 has had a stronger positive impact on the Aymara than on other ethnic groups. In addition, we have suggested that the differences in outcomes are linked to differences in preferences, and therefore probably to differences in behavior of the beneficiaries from different ethnic groups. It is an interesting question for further research whether our finding that the poor older Peruvians with an Aymara background give a higher weight to health can be confirmed in other contexts (for instance in Bolivia, the country with the largest Aymara subpopulation).

Our findings support the growing attention of policy makers to the specific situation of indigenous people and may improve the targeting effectiveness of social policies. Interpreting ethnic preference differences in terms of the different weights attached to various dimensions of life is just a modest first step towards the recognition of different ethnic identities in the evaluation of policy. It remains still far removed from a more ambitious story about heterogeneity in world views and culturally inspired ideas about development (see, for instance, Sen (1999, 2004)). Self-determination and participation in policy design can and should go much further. Moreover, preferences of indigenous groups may change when confronted with other ways of life, e.g., by moving from a traditional rural to a modern urban environment.

A comparison of the findings obtained through quantitative surveys like ours to the outcomes of well-structured deliberative processes with indigenous groups would certainly be an interesting avenue to pursue. Such a comparison could also offer new insights in the realism of the existence of a stable and complete preference relation, that has been assumed implicitly in our empirical work. The interest in exploring the consequences of incompleteness of the preference relation goes far beyond ethnic differences and is an essential step if one wants to incorporate insights from behavioral economics into policy evaluation (Bernheim (2009); Bernheim and Rangel (2009); Fleurbaey and Schokkaert (2013); Decancq and Nys (2018)).

At a more basic level, we believe that it is necessary to develop better methods of measuring individual preferences for non-market commodities (such as health and respect). In addition to methods based on revealed preferences and life

satisfaction regressions, attention should also be given to contingent valuation and other stated preferences techniques. Better identification of preferences for non-market commodities is a *conditio sine qua non* for the development of methods for policy evaluation that are based on the ordinal preferences of the people involved.

# References


Adler, M. D., P. Dolan, and G. Kavetsos (2017): "Would you choose to be happy? Tradeoffs between happiness and the other dimensions of life in a large population survey," *Journal of Economic Behavior and Organization*, 139, 60 – 73.

Adler, M. D. and M. Fleurbaey, eds. (2016): *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press.

Artaraz, K. and M. Calestani (2015): "Suma qamana in Bolivia: Indigenous Understandings of Well-being and their Contribution to a Post-neoliberal Paradigm," *Latin American Perspectives*, 42, 216–233.

Bando, R., S. Galiani, and P. Gertler (2016): "The Effects of Non-Contributory Pensions on Material and Subjective Well Being," Tech. Rep. 22995, National Bureau of Economic Research.

Beegle, K., K. Himelein, and M. Ravallion (2012): "Frame-of-reference bias in subjective welfare," *Journal of Economic Behavior and Organization*, 81, 556–570.

Benjamin, D. J., O. Heffetz, M. S. Kimball, and A. Rees-Jones (2012): "What Do You Think Would Make You Happier? What Do You Think You Would Choose?" *American Economic Review*, 102, 2083–2110.

Bernal, N., M. A. Carpio, and T. Klein (2017): "The effects of access to health insurance: evidence from a regression discontinuiy design in Peru," *Journal of Public Economics*, 154, 122–136.

Bernheim, B. D. (2009): "Behavioral welfare economics," *Journal of the European Economic Association*, 7, 267–319.

Bernheim, B. D. and A. Rangel (2009): "Beyond revealed preference: choice theoretic foundations for behavioral welfare economics," *Quarterly Journal of Economics*, 124, 51–104.

Bosmans, K., K. Decancq, and E. Ooghe (2018): "Who is affraid of aggregating money metrics?" *Theoretical Economics*, 13, 467–484.

Calestani, M. (2009): "An Anthropology of 'The Good Life' in the Bolivian Plateau," *Social Indicators Research*, 90, 141–153.

Camacho, A. and E. Conover (2011): "Manipulation of Social Program Eligibility," *American Economic Journal: Economic Policy*, 41–65.

Canavire Bacarreza, G. J., A. Chong, F. Rios-Avila, and M. Yanez Pagans (2017): "Will elders provide for their grandchildren?: unconditional cash transfers and educational expenditures in Bolivia," Tech. rep., World Bank Policy Research Working Paper 8176.

Clark, A. E. and A. J. Oswald (2002): "A simple statistical method for measuring how life events affect happiness," *International Journal of Epidemiology*, 31, 1139–1144.

Copestake, J., M. Guillen-Royo, W.-J. Chou, T. Hinks, and J. Velazco (2009): "The relationship between economic and subjective well-being indicators in Peru," *Applied Research in Quality of Life*, 4, 155–177.

Deaton, A. (1979): "The distance function in consumer behaviour with applications to index numbers and optimal taxation," *Review of Economic Studies*, 46, 391–405.

Deaton, A. and J. Muellbauer (1980): *Economics and Consumer Behavior*, Cambridge University Press.

Decancq, K., M. Fleurbaey, and E. Schokkaert (2015a): "Happiness, equivalent incomes, and respect for individual preferences," *Economica*, 82, 1082–1106.

——— (2015b): "Inequality, Income, and Well-Being," in *Handbook of Income Distribution*, ed. by A. B. Atkinson and F. Bourguignon, Elsevier, vol. 2 of *Handbook of Income Distribution*, 67 – 140.

——— (2017): "Well-being inequality and preference heterogeneity," *Economica*, 84, 210–238.

Decancq, K. and A. Michiels (2018): "Measuring Successful Aging With Respect for Preferences of Older Persons," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, (forthcoming).

Decancq, K. and D. Neumann (2016): "Does the choice of well-being measure matter empirically? An illustration with German data," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler and M. Fleurbaey, Oxford University Press.

Decancq, K. and A. Nys (2018): "Non-parametric wellbeing comparisons," *CORE Discussion Paper 2018/21*.

Decancq, K. and E. Schokkaert (2015): "Beyond GDP: Using Equivalent Incomes to Measure Well-Being in Europe," *Social Indicators Research*, 1–35.

Ferreira, F. H. G. and M. A. Lugo (2013): "Multidimensional Poverty Analysis: Looking for a Middle Ground," *The World Bank Research Observer*, 28, 220–235.

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

Fleurbaey, M. and D. Blanchet (2013): *Beyond GDP: Measuring Welfare and Assessing Sustainability*, Oxford University Press.

Fleurbaey, M. and F. Maniquet (2011): *A Theory of Fairness and Social Welfare*, Cambridge University Press.

Fleurbaey, M. and E. Schokkaert (2013): "Behavioral Welfare Economics and Redistribution," *American Economic Journal: Microeconomics*, 5, 180–205.

Fleurbaey, M. and K. Tadenuma (2014): "Universal Social Orderings: An Integrated Theory of Policy Evaluation, Inter-Society Comparisons, and Interpersonal Comparisons," *Review of Economic Studies*.

Forbes, D. (1870): "On the Aymara indians of Bolivia and Peru," *The Journal of the Ethnological Society of London (1869-1870)*, 2, 193–305.

Fujiwara, D. and P. Dolan (2016): "Happiness-based policy analysis," in *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, 286–320.

Galiani, S., P. Gertler, and R. Bando (2016): "Non-contributory pensions," *Labour Economics*, 38, 47 – 58.

Gana, K., N. Bailly, Y. Saada, M. Joulain, and D. Alaphilippe (2012): "Does life satisfaction change in old age: Results from an 8-year longitudinal study," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 68, 540–552.

GuillenRoyo, M. (2008): "Consumption and subjective well-being: exploring basic needs, social comparison, social integration and hedonism in Peru," *Social Indicators Research*, 89, 545–555.

——— (2011): "Reference group consumption and the subjective well-being of the poor in Peru," *Journal of Economic Psychology*, 32, 259–272.

Kahneman, D. and A. B. Krueger (2006): "Developments in the measurements of subjective well-being," *Journal of Economic Perspectives*, 20, 3–24.

Kant, S., I. Vertinsky, B. Zheng, and P. M. Smith (2014): "Multi-Domain Subjective Wellbeing of Two Canadian First Nation Communities," *World Development*, 140–157.

Moreno, M. (2014): "Patrones de autoidentificación etnorracial de la población indígena en las encuestas de hogares en el Perú," *Debates en Sociología*, 39–71.

Neelsen, S. and O. O'Donnell (2017): "Progressive universalism? The impact of targeted coverage on health care access and expenditures in Peru," *Health Economics*.

Ñopo, H., J. Saavedra, and M. Torero (2004): "Ethnicity and Earnings in Urban Peru," Tech. Rep. 980, IZA Discussion Paper Series.

Olivera, J. and I. Tournier (2016): "Succesful ageing and poverty: the case of Peru," *Ageing & Society*, 36, 1690–1714.

Paredes, M. (2007): "Fluid identities: exploring ethnicity in Peru," Tech. Rep. 40, CRISE Working Paper, University of Oxford.

Pasquier-Doumer, L. and F. Risso Brandon (2015): "Aspiration failure: a poverty trap for indigenous children in Peru," *World Development*, 72, 208–223.

Ravallion, M., K. Himelein, and K. Beegle (2016): "Can subjective questions on economic welfare be trusted?" *Economic Development and Cultural Change*, 64, 697–726.

Samuelson, P. A. (1974): "Complementarity: an essay on the 40th anniversary of the Hicks-Allen revolution in demand theory," *Journal of Economic Literature*, 12, pp. 1255–1289.

——— (1977): "Reaffirming the existence of" reasonable" bergson-samuelson social welfare functions," *Economica*, 81–88.

Sen, A. K. (1985): *Commodities and Capabilities*, Amsterdam and Oxford: North-Holland.

——— (1999): *Development as Freedom*, Oxford: Oxford University Press.

——— (2004): "Capabilities, lists, and public reason: continuing the conversation," *Feminist Economics*, 10, 77–80.

Smith, T. (2006): *With One Heart: The Panao Quechua Way of Life*, Lima: Summer Institute of Linguistics.

Stone, A. A., J. E. Schwartz, J. E. Broderick, and A. Deaton (2010): "A snapshot of the age distribution of psychological well-being in the United States," *Proceedings of the National Academy of Sciences*, 107, 9985–9990.

Sulmont, D. (2011): "Race, ethnicity and politics in three Peruvian localities: An analysis of the 2005 CRISE Perceptions Survey in Peru," *Latin American and Caribbean Ethnic Studies*, 6, 47–78.

Van de gaer, D., J. Vandenbossche, and J. L. Figueroa (2013): "Children's health opportunities and project evaluation: Mexico's Oportunidades program," *The World Bank Economic Review*, 28, 282–310.

Van Praag, B. and B. Baarsma (2005): "Using happiness surveys to value intangibles: the case of airport noise," *Economic Journal*, 115, 224–246.

World Bank (2015): *Indigenous Latin America in the Twenty-First Century: the First Decade*, Washington D.C.: World Bank.

Zweifel, P., F. Breyer, and M. Kifmann (2009): *Health Economics*, Berlin Heidelberg: Springer.

# Appendix. Additional Material

## Sample selection

We construct the final sample as follows (see Table 12). We start from an initial sample of 8,089 observations. First, we drop 349 observations who have received the program without being eligible.[32] Eligibility conditions require that recipients of the Pension 65 program should not be pensioners, nor be affiliated to any contributory pension system. These conditions are assumed to apply for the treatment group, but in order to have a control group which is as similar as possible, we also drop 252 respondents from the control group who do not comply with these eligibility criteria. Second, we drop 356 observations with missing values in relevant variables. This gives us an intermediate sample of 7,384 observations: 3,869 respondents in the 2012 wave and 3,515 in the follow-up wave. Third, 932 respondents were only present in one wave, and were dropped from the sample. Finally, we restrict our sample to the respondents who self-identify to belong to one of the three largest ethnic groups in Peru: Mestizo, Quechua, and Aymara. That means that 2,516 respondents with a fluid ethnicity were dropped. The final (balanced) sample consists of 1,968 respondents in both waves.

| | Total | 2012 | 2015 |
|---|---|---|---|
| Initial sample | 8,089 | 4,242 | 3,847 |
| - Non-eligible | 349 | 219 | 130 |
| - Missing data | 356 | 154 | 202 |
| Intermediate sample | 7,384 | 3,869 | 3,515 |
| - Attrition from panel | 932 | 643 | 289 |
| - Fluid ethnicity | 2,516 | 1,258 | 1,258 |
| Final sample | 3,936 | 1,968 | 1,968 |

Table 12: Sample selection

Table 13 focuses on the 3,869 respondents of the intermediate sample in the 2012 wave by means of three separate logit regressions. The first regression (column (1)) focuses on the respondents who are not present in the follow-up wave. The second column shows the characteristics of respondents with a fluid identity, i.e.,

[32] There are 86 respondents who received the transfer in the baseline survey in 2012 and for 3 respondents there was no information on whether they were receiving the transfer or not, and 8 respondents were younger than 65.

respondents who report a different ethnic background in both waves. The third column shows respondents who are present in both waves and show a consistent identity. Older, more unhealthy and less satisfied respondents have a lower probability of being present in both waves. Male respondents are more likely to show a consistent identity. Respondents who report an indigenous background in the first wave, in particular those who identify as Aymara, are found to have a more consistent identity compared to respondents with a Mestizo background.

## Detailed health variables

Table 14 presents the effect of the introduction of the Pension 65 program on some additional health variables.

| | Attrition from sample (1) | Fluid identity (2) | Consistent identity (3) |
|---|---|---|---|
| Expenditures | 0.020 | -0.216*** | 0.177** |
| | (0.092) | (0.075) | (0.071) |
| Health | -0.005* | 0.002 | 0.001 |
| | (0.003) | (0.002) | (0.002) |
| ADL | -0.006*** | 0.003 | 0.002 |
| | (0.002) | (0.002) | (0.002) |
| Respect | -0.002 | 0.002 | -0.001 |
| | (0.002) | (0.002) | (0.002) |
| SWB | -0.014*** | -0.000 | 0.008*** |
| | (0.004) | (0.003) | (0.003) |
| Altitude | -0.028 | 0.079** | -0.057* |
| | (0.035) | (0.032) | (0.029) |
| Urban | 0.204* | 0.031 | -0.145* |
| | (0.110) | (0.089) | (0.084) |
| Age | 0.054*** | -0.013 | -0.020** |
| | (0.010) | (0.008) | (0.008) |
| Male | 0.113 | -0.522*** | 0.382*** |
| | (0.097) | (0.078) | (0.073) |
| Marital status | -0.070 | -0.074 | 0.113 |
| | (0.099) | (0.080) | (0.077) |
| Work status | -0.004*** | 0.001 | 0.002** |
| | (0.001) | (0.001) | (0.001) |
| Quechua | -0.284** | -1.155*** | 1.117*** |
| | (0.120) | (0.108) | (0.093) |
| Aymara | -0.307 | -2.355*** | 1.766*** |
| | (0.219) | (0.304) | (0.190) |
| Constant | -3.594*** | 0.877 | -0.352 |
| | (1.009) | (0.815) | (0.775) |
| Observations | 3869 | 3869 | 3869 |
| Pseudo $R^2$ | 0.048 | 0.061 | 0.060 |

Robust standard errors in parentheses. * $< 0.10$, ** $< 0.05$ , *** $< 0.01$

Table 13: Sample selection

| | Δ Physical | Δ Physical role | Δ Bodily pain | Δ General health | Δ Health better | Δ MNA score |
|---|---|---|---|---|---|---|
| Treatment | 0.565 | 3.356 | 3.162 | -0.038 | 0.005 | 0.260* |
| | (1.493) | (3.037) | (2.296) | (1.290) | (0.036) | (0.136) |
| Treatment × Quechua | 1.364 | 1.623 | -1.927 | 2.307 | 0.087 | 0.049 |
| | (2.770) | (5.213) | (4.223) | (2.439) | (0.066) | (0.260) |
| Treatment × Aymara | 7.361 | 1.509 | 13.890** | 7.716* | 0.115 | 0.887* |
| | (4.963) | (8.690) | (6.817) | (3.962) | (0.127) | (0.472) |
| Δ Altitude | -0.663 | -4.699 | 2.709 | 4.115 | -0.104** | -0.285 |
| | (3.336) | (3.750) | (3.772) | (2.824) | (0.050) | (0.311) |
| Δ Urban status | -11.233** | -7.169 | 4.214 | 1.875 | -0.012 | -0.256 |
| | (5.388) | (11.891) | (8.824) | (5.774) | (0.150) | (0.519) |
| Δ Marrital status | -1.250 | 12.195*** | -4.558 | -2.125 | -0.049 | -0.108 |
| | (2.570) | (4.504) | (4.200) | (2.287) | (0.062) | (0.268) |
| Δ Work status | 0.086*** | 0.113*** | 0.041** | 0.023** | 0.000 | 0.005*** |
| | (0.012) | (0.023) | (0.018) | (0.011) | (0.000) | (0.001) |
| Constant | -3.671*** | -2.704 | 0.952 | 0.159 | -0.024 | -0.020 |
| | (1.020) | (2.043) | (1.584) | (0.877) | (0.025) | (0.098) |
| Constant × Quechua | -3.062* | 5.808* | -3.183 | -3.486** | -0.067 | -0.472*** |
| | (1.814) | (3.434) | (2.743) | (1.569) | (0.043) | (0.181) |
| Constant × Aymara | 8.217*** | 11.614** | 1.899 | -1.895 | 0.058 | 0.201 |
| | (3.042) | (4.998) | (4.335) | (2.533) | (0.078) | (0.307) |
| Observations | 1968 | 1968 | 1968 | 1968 | 1901 | 1938 |
| $R^2$ | 0.047 | 0.023 | 0.013 | 0.011 | 0.007 | 0.029 |

Robust standard errors in parentheses. * $< 0.10$, ** $< 0.05$ , *** $< 0.01$

Table 14: Difference-in-differences estimators for different health aspects