Ethnic differences and preference heterogeneity: assessing social pensions in Peru

Koen Decancq, Javier Olivera & Erik Schokkaert

Published online: 29 Dec 2024.

Submit your article to this journal ⃝

Article views: 115

View related articles ⃝

View Crossmark data ⃝

Routledge
Taylor & Francis Group

Ethnic differences and preference heterogeneity: assessing social pensions in Peru

Koen Decancq[a,b,c], Javier Olivera[d,e,f] and Erik Schokkaert[b]

[a]Centre for Social Policy, University of Antwerp, Antwerp, Belgium; [b]Department of Economics, KU Leuven, Leuven, Belgium; [c]Centre for Philosophy of Natural and Social Science, London School of Economics, London, UK; [d]Luxembourg Institute of Socio-Economic Research (LISER), Belval, Luxembourg; [e]Department of Economics, Pontificia Universidad Catolica del Peru, Lima, Peru; [f]National Bank of Belgium, Brussels, Belgium

ABSTRACT

We examine the impact of the non-contributory social pension program (Pension 65) in Peru, highlighting its varying effects on the three main ethnic groups: Mestizo, Quechua, and Aymara. Notably, Aymara beneficiaries have experienced greater improvements in health outcomes compared to other Peruvians. To account for these ethnic differences when evaluating policy programs, it is essential to use a welfare criterion that reflects preference heterogeneity. We propose a natural criterion: a program benefits a recipient if it lifts them to a higher indifference curve. We contrast this approach with an alternative that uses self-reported subjective well-being to evaluate a policy program. Through a panel life satisfaction regression, we find evidence of preference heterogeneity between the Aymara and other ethnic groups, consistent with the observed differences. Lastly, we explore why, contrary to simple intuition, not all beneficiaries reach a higher indifference curve.

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, this principle seems to have had only a minimal impact on the concrete practice of policy evaluation.

The standard approach to evaluate a policy 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, or on non-monetary variables such as health or educational outcomes. Yet, when a policy program affects the multiple dimensions of well-being differently, a dimension-by-dimension approach does not provide an overall picture of well-being. It remains blind to the phenomenon of cumulative deprivation, which arises when the same individuals are deprived in different dimensions of life (Decancq 2023; Ferreira and Lugo 2013). A multidimensional welfare criterion that integrates and weighs the effects on the different dimensions is therefore needed. In this paper we address the question how the preferences of the individual can be used to construct such a multidimensional welfare criterion and how it is influenced by preference heterogeneity across ethnic groups.

Building on the literature in welfare economics and social choice theory following Arrow (1951), we assume that individuals have preferences over various aspects of their lives, including non-monetary dimensions. These preferences are used to evaluate their life situation in a well-considered and detached manner or to compare it with the situations of others. It captures their views on the 'good life'. Naturally, such preferences may evolve over time and vary across individuals or ethnic groups, capturing their different views on the 'good life'. Moreover, these preferences are not always revealed by actual choice behavior, which may be shaped by constraints, informational imperfections, or imperfect decision-making heuristics.

A possible candidate as welfare criterion is self-reported subjective well-being. However, while a subjective well-being measure may be consistent with individual preferences, it is also sensitive to changes in aspirations, expectations, and reference frames. To overcome this unattractive feature, we will propose an alternative approach that is solely based on the ordinal preferences of the different ethnic groups as captured by the shape of their indifference curves. We say that a program benefits a recipient if they are lifted to a higher indifference curve according to their own preferences. This criterion uses only ordinal intrapersonal well-being comparisons and does not involve any debatable normative choices that are needed to cardinalize utility in an interpersonally comparable way. We discuss these two approaches in Section 2.

We illustrate the empirical relevance of the choice of an adequate normative criterion through an analysis of the Pension 65 program, a non-contributory pension program in Peru. Like other Latin American countries, Peru is characterized by a large socio-economic gap between indigenous and non-indigenous people (World Bank 2015). We compare the effects of the policy program on the two largest indigenous groups, the Quechua and the Aymara, with the effects on the group of Peruvians who identify themselves as having 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 individuals received a transfer from the program (in 2012 and 2015). In Section 3 we give some background about the Pension 65 program and about the measurement of ethnicity in Peru. In Section 4, we compare the beneficiaries of the Pension 65 program to a control group of Peruvians who are classified as poor, but not as extremely poor and, hence, were not eligible for the program. We show that the Pension 65 program has had different effects on the three ethnic groups and, most strikingly, that the Aymara beneficiaries of the program have experienced larger increases in health compared to other Peruvians.

To apply the proposed welfare criterion, we need information about the preferences of the different ethnic groups. Following Decancq, Fleurbaey, and Schokkaert (2015a, 2015b, 2017), we derive this information 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). Our analysis of preferences is presented in Section 5. We then apply our preference-based criterion for policy evaluation in Section 6, and we show that taking into account ethnic differences in the conception of a good life is indeed relevant for the evaluation of the Pension 65 program. We find that not all beneficiaries have been lifted to a higher indifference curve. At first sight, this may seem surprising, since receiving a pension should have broadened their opportunity sets. However, our welfare criteria are at the level of the individual, while the pension can be used at the household level to invest or improve the situation of other household members and may also have had an effect on other transfers to and from the household (Olivera and Iparraguirre 2024).

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, Carpio, and Klein (2017) and Neelsen and O'Donnell (2017) investigate 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 and without analysing the ethnic differences. Olivera and Tournier (2016) use the baseline wave of the ESBAM data to analyse 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, Galiani, and Gertler (2020). They use a different econometric strategy to identify the effects of the Pension 65 program and analyse the impact of the program on different life dimensions separately without looking at an overall well-being measure. Bernal, Olivera, and Suhrcke (2024) and Valderrama and Olivera (2023) show that the Pension 65 program has had positive effects on health and on life expectancy. None of the papers on Peru have explicitly analysed 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, Vandenbossche, and Figueroa (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.

2. Two policy evaluation criteria

When a policy program affects different life dimensions, an aggregation of these different effects is necessary to arrive at an overall evaluation of individual well-being. The weights used in this aggregation procedure can be set in different ways (see Decancq and Lugo 2013 for an overview), but the most popular approach in economics is to base them on the preferences of the individuals concerned. How to turn this general principle in an operational approach is not straightforward, however, *a fortiori* when different individuals or sub-groups have different preferences. Ethnic differences in preferences are a particularly striking and policy-relevant example. In this section, we discuss two welfare criteria for program evaluation that are sensitive to preferences. These criteria are inspired by the recent literature on well-being measurement.[2]

A first possibility is to focus on a measure of subjective well-being. This leads to the following welfare criterion for policy evaluation:

Subjective well-being criterion *A policy program benefits a recipient if they report a higher subjective well-being after the introduction of the policy program.*

This criterion has appealing features. It is sensitive to preference differences between ethnic groups or individuals in general, if the subjective well-being measure is consistent with the preferences of the respondents. This consistency assumption requires that respondents at a given point of time report a higher subjective well-being index in situations that they prefer (which reflect a better life according to them). This consistency question seems plausible, but is hard to test empirically.

However, the subjective well-being criterion has also some normative drawbacks. The most important is that it does not only capture the outcomes in the different dimensions of life as weighted by preferences, but also the differences that are due to changing aspirations and expectations, or to mere individual idiosyncrasies in the use of the response scale.[3] Subjective well-being is a specific cardinalization of the indifference map. Figure 1 illustrates for the case of two life dimensions (expenditures and health). Suppose a policy program moves a person from situation A to situation B. They have moved to a higher indifference curve and prefer the situation after the introduction to the situation before the introduction of the policy program. Yet, it is possible that their subjective well-being is lower in B
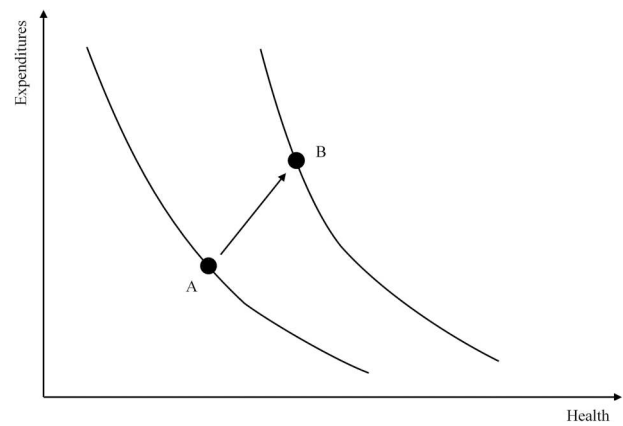
Figure 1. Preferences and life satisfaction.

compared to A. They can be confronted with other people that do even better than them, for instance, or maybe something miserable has happened in their personal life. These events may have shifted their frame of reference so that they report a lower subjective well-being, despite the fact that they moved to a higher indifference curve. While they prefer the situation obtained through the policy program both before and after its implementation, the change in their aspirations may lead them to report a lower level of subjective well-being after the policy program.

Ultimately, it is a normative question whether we want to take up differences in aspirations and reference frames when evaluating a policy program. If one thinks that the objective situation as evaluated by the ordinal preferences is what should matter for policy evaluation, rather than subjective experiences, the subjective well-being criterion is not attractive. The following preference-based criterion presents a natural alternative.

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

This criterion is indeed very natural, to the extent that it is almost tautological for economists, at least if the additional pension benefit is used to improve the well-being of the individual beneficiary. It is consistent with all well-being measures that are based on the information of ordinal preferences alone, such as the equivalent income measure, the quantity-metric, or money-metric well-being measures as discussed by Deaton and Muellbauer (1980, 179–182).[4] Essentially, these well-being measures 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 assign labels to indifference curves, we take a

more robust approach here and limit ourselves to ordinal preference information captured by the shape of the indifference curves. This robustness comes at a price, however. Our proposed criterion is able to tell us whether the program was beneficial for an individual or not, but it does not provide us with an exact quantification of how much the program has benefited the person at hand. By focusing on each individual's situation separately, the criterion avoids the need for interpersonal comparisons.

3. Pension policy in Peru: some background information

We will illustrate the use of the two welfare criteria for the evaluation of the introduction of the Pension 65 program in Peru. The program is described in Section 3.1. Peru is an ethnically heterogeneous country and we want to investigate whether the ethnic differences matter for the evaluation of this policy program. We discuss how to measure ethnicity in Peru in Section 3.2.

3.1. 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 roll-out of the program started in the poorest districts of 6 prioritized regions and then was rapidly extended to all districts, reaching about 500,000 recipients by 2015. It is the second largest social program in Peru, behind the conditional cash transfer program *Juntos*, covering about 20% of the population aged 65 or over, and costing 0.11% of GDP.

The program has two components. First, recipients receive 250 Peruvian soles (about US$76 in 2015) every two months. This transfer amounts to about 18% of the total expenditures of the targeted group before the intervention. In addition, the program facilitates registration in the public health system (Seguro Integral de Salud, SIS), which covers health at no cost, although it can incur some out-of-pocket expenditures.[5] 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 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.[6] Based

on a comparison of their SISFOH score with region-specific 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 are determined independently from the thresholds, which avoids the possibility of manipulation (see Camacho and Conover 2011). Indeed, Bando, Galiani, and Gertler (2020) illustrate that there is no bunching of the SISFOH scores around the eligibility cut-off, which indicates that manipulation is unlikely. Additionally, Bernal, Olivera, and Suhrcke (2024) and Valderrama and Olivera (2023) present statistical evidence showing no manipulation of the SISFOH score in the ESBAM sample.

3.2. Measuring ethnicity in Peru

The complex interplay between ethnicity, exclusion, segregation, and race makes it difficult to define and measure ethnicity in quantitative studies.[7] 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.[8] 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.[9] 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).[10] 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 background) to avoid the stigma of being non-white (Ñopo, Saavedra, and Torero 2004; Pasquier-Doumer and Brandon 2015). As a consequence, ethnic identification can be influenced by the relative numbers of indigenous people in the localities (Moreno 2014; Sulmont 2011). 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 most obvious 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.[11] 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).

4. The dimension-by-dimension impact of the Pension 65 program

We first describe the observational sample that we use to estimate the impact of the Pension 65 program (Section 4.1) and the outcomes of the program for the different ethnic groups (Section 4.2). We then analyse these outcomes in more detail to gage the impact of the policy program (Section 4.3).

4.1. Sample

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.[12] 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 had been updated at

the moment of the sampling (unshaded districts in Figure 2 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 within 0.3 standard deviations above or below the SISFOH threshold for extreme poverty. The goal of this design 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 condition itself. We say that individuals living in households located below or above the eligibility threshold fall into the treatment and control group, respectively (for more details, see the methodological note by MIDIS 2013).

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 4242 individuals (in 3194 households) and there are 3847 individuals (in 2967 households) in the follow-up wave of 2015.

We restrict the sample to the respondents who 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.[13] Our final sample is composed of 1775 respondents who are observed in both waves, from which 1122 (63%), 500 (28%) and 153 (9%) respondents are Mestizo, Quechua and Aymara, respectively. Clearly, this sample is not
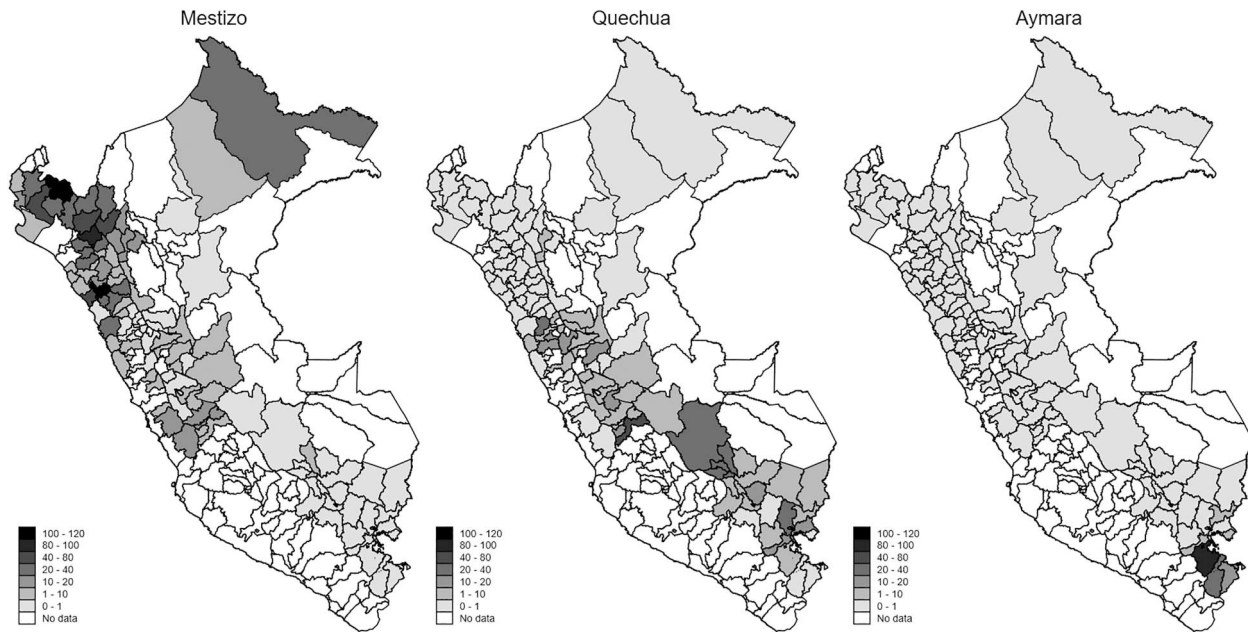
Figure 2. Ethnicity by region.

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. Table 1 shows the distribution of ethnicity and recipient condition in our sample. The information about whether the individual is a recipient of the transfer comes from the administrative records of the program. We use the recipient condition to identify our control and treatment groups. Out of 1146 eligible individuals in our sample, 1032 actually received the transfer, and 114 did not receive it. Regarding the ineligible individuals of our sample (629), 622 did not receive the transfer, but only 7 received it. Thus, there is a high correlation (0.86) between being eligible to the program and being a recipient.

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 2, 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 meter above sea-level), compared to the Quechua (3250 m) and Mestizo (1750 m). 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 residence area effects and altitude of the district in which the respondent lives.

4.2. Summary statistics

Let us start by looking at four dimensions of life that will play a central role in our analysis: expenditures, health, activities of daily living (ADL), and respect. Table 2 presents averages for 2012 and 2015 broken down by ethnic group and treatment status. We also show the outcomes for subjective well-being.

Information about *household expenditures* is collected for eight types of goods, by source of acquisition (buying, gift, and self-consumption).[14] Expenditures on food make up a large share of household expenditures (about 52%) and a high proportion of expenditure corresponds to self-consumption (about 21%). 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. This method allows us to correct for household needs, but does not take into account the distribution of consumption within the household. Between both waves, the average household size in our sample is reduced from 3.03 to 2.86 persons (with slightly smaller households for the Aymara subpopulation). As is clear from Table 2, expenditures increase, on average, for the

Table 1. Control and treatment groups by ethnicity.

	Mestizo	Quechua	Aymara	Total
Control	407	241	88	736
Treatment	715	259	65	1039
Total	1122	500	153	1775

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

		Total	Mestizo		Quechua		Aymara	
			C	T	C	T	C	T
2012	Expenditure	5.75	5.88	5.76	5.78	5.67	5.53	5.48
	Health	50.02	53.79	49.75	49.36	46.93	49.62	44.71
	ADL	79.78	84.40	80.79	77.00	75.51	73.61	75.42
	Respect	84.14	86.77	87.95	82.09	78.47	68.59	76.96
	SWB	53.43	57.02	56.64	49.04	49.04	44.08	42.05
2015	Expenditure	5.80	5.93	5.83	5.80	5.84	5.16	5.46
	Health	48.69	50.92	48.68	46.03	45.24	50.72	55.64
	ADL	75.77	79.53	78.76	69.99	66.39	74.55	79.91
	Respect	82.81	84.40	86.09	77.80	76.27	79.24	86.22
	SWB	52.92	54.55	55.02	49.85	49.27	48.20	52.01

Notes: Expenditure is expressed in logs, and the other variables are indices ranging from 0 to 100.

treated Mestizo and Quechua households, but not for the Aymara households.

A *health* index is constructed based on the first (polychoric) principal component of a battery of four 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.[15] 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.

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 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. As we have discussed in Section 2, 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.

4.3. 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. For this purpose, 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.[16] To be precise, let

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

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

$$\Delta o_i = \beta \triangle z_i + \gamma + \eta P65_i + \triangle \varepsilon_i. \quad (2)$$

In addition, we will interact the treatment parameter η and the time trend γ with ethnicity dummies. This

specification allows for heterogeneous treatment effects and different time trends across the ethnic groups in the control group.

The treatment parameter(s) 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−value = 0.000). As we can expect the SISFOH score to be correlated with the expenditures, this result is not surprising. Another significant difference is found with respect to urban status. This was also to be expected as the sampling framework captures the fact that extreme poor households tend to be more concentrated in rural areas.[17] An alternative method to inspect the balancedness of treatment and control group in the pre-treatment period is the normalized difference approach of Imbens and Rubin (2015). The normalized difference is a scale-free measure of the difference in locations, and is defined as the difference in means between the treatment and control groups, divided by the square root of half the sum of the treatment and control group variances. The normalized differences reported in the last column of Table 3 confirm that control and treatment groups are not perfectly balanced (mainly for expenditures, health, gender and urban status) but differences of 25% or less are not very problematic in the light of the rules

of thumb that are suggested by Imbens and Rubin (2015).

The treatment parameter η would capture 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. The bias that can occur due to pre-treatment differences between the control and treatment groups is mitigated, but not solved, by our introduction of heterogeneous time trends for the different ethnicities. We therefore consider our results as illustrative only. They illustrate the relevance of the choice of a normatively attractive measure of well-being, but sufficient care is needed in interpreting our estimates of η as the causal effect of the policy program.

Table 4 shows the estimates of equation (2) for the four 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 ethnicity dummies on the one hand and the treatment dummy and the time trend (i.e. the constant in our estimation in first differences) on the other hand. 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 2), 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.[18] 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, ADL and respect, 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 program recipient does have a positive effect (of about 10%) on equivalized expenditures (see column (1) of Table 4). Yet, when taking into account interactions (column (2) of Table 4), we see that the pattern for the Aymara differs remarkably. There is a significant decrease in expenditures for the Aymara in the control group, such that, compared to

Table 3. Balancedness of variables in 2012.

	Control group	Treatment group	t-Value	p-Value	Normalized difference
Expenditure	5.80	5.72	−3.5	0.000	−16.9%
Health	51.84	48.74	−3.26	0.001	−15.7%
ADL	80.68	79.14	−1.33	0.183	−6.4%
Respect	83.07	84.90	1.7	0.089	8.2%
Altitude	7.27	7.41	2.36	0.018	11.2%
Urban	0.44	0.30	−6.1	0.000	−29.2%
Married	0.71	0.73	1.29	0.197	6.2%
Working	0.74	0.76	1.08	0.279	5.2%
Age	71.30	71.35	0.22	0.824	1.1%
Male	0.55	0.63	3.39	0.001	16.3%

Notes: Expenditure is expressed in logs; health, ADL and respect are indices ranging from 0 to 100; the altitude (meters above the sea level) of the district is expressed in logs; urban, married, working, and male are expressed in shares. The normalized difference is the difference between the treatment and control group, divided by the square root of half the sum of the treatment and control group variances, see Imbens and Rubin (2015).

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

	Δ Expenditures		Δ Health		Δ ADL		Δ Respect	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.102***	0.020	2.303**	2.132	1.815	3.089*	0.201	0.333
	(0.029)	(0.036)	(1.058)	(1.306)	(1.448)	(1.773)	(1.467)	(1.733)
Treatment × Quechua		0.136**		−0.462		−4.747		2.173
		(0.062)		(2.468)		(3.331)		(3.385)
Treatment × Aymara		0.345***		7.860**		0.907		−1.668
		(0.111)		(3.961)		(5.335)		(6.120)
Δ Altitude	−0.109**	−0.120**	1.697	1.695	−6.642**	−6.321*	−6.225***	−6.226***
	(0.049)	(0.055)	(2.740)	(2.767)	(3.293)	(3.262)	(2.042)	(2.124)
Δ Urban status	0.423***	0.403***	−0.161	0.228	−10.102	−9.030	−16.712***	−16.502***
	(0.114)	(0.119)	(5.136)	(5.116)	(6.184)	(6.273)	(5.581)	(5.674)
Δ Marital status	−0.078	−0.075	0.273	0.339	2.759	2.440	2.678	2.773
	(0.061)	(0.061)	(2.387)	(2.381)	(3.357)	(3.374)	(3.388)	(3.327)
Δ Work status	0.000	0.000	0.066***	0.066***	0.060***	0.061***	−0.023*	−0.024*
	(0.000)	(0.000)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)	(0.014)
Constant	−0.012	0.051*	−1.782**	−2.167**	−4.266***	−4.234***	−1.648	−2.500*
	(0.022)	(0.028)	(0.795)	(1.009)	(1.092)	(1.363)	(1.173)	(1.420)
Constant × Quechua		−0.037		−0.305		−2.190		−2.111
		(0.046)		(1.784)		(2.431)		(2.518)
Constant × Aymara		−0.430***		4.030		5.676		12.896***
		(0.077)		(2.490)		(3.791)		(4.390)
Observations	1775	1775	1775	1775	1775	1775	1775	1775
R^2	0.022	0.047	0.026	0.037	0.014	0.025	0.006	0.021

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

this refined control group, the treatment effect for the treated Aymara becomes 0.365, which is significantly larger than the effect for the Mestizos and the Quechua (p−value = 0.0005). Apparently, in contrast with 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.[19]

Even more striking is the treatment effect on health: it is positive, but the simple model without interactions of column (3) masks an interesting heterogeneous effect across ethnic groups. In fact, we see in column (4) that the effect of the treatment on health is large and significant for the Aymara (p−value = 0.0076), but not significantly positive for the other ethnic groups.[20] In fact, we can notice that the health of the Aymara in the control group has also increased, albeit at a much lower rate

than for the Aymara in the treatment group. Interethnic heterogeneity is much less for ADL (see columns (5) and (6) of Table 4).

In columns (7) and (8) of Table 4, we see that for the respect variable the treatment effects are not precisely estimated. Respect strongly 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).

All-in-all, our analysis confirms that introducing ethnic differences reveals interesting heterogeneous effects of the program that would otherwise have remained hidden.[21] In Section 6 we will return to the normative implications of these findings.
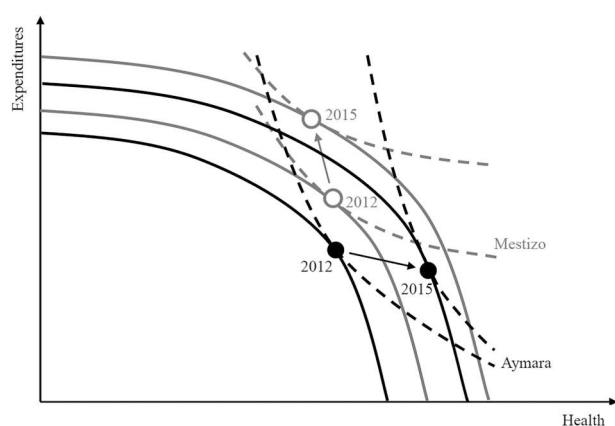
Figure 3. Heterogeneous effects of a program when preferences differ.

5. The Aymara: a different view of the good life?

The findings in the previous section show that Pension 65 seems to have a strong 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. The standard economic approach would be to focus on opportunities and preferences to explain these differences. This analysis is at the level of the individual, while the effects of the Pension 65 program are located at the household level. Still, with this caveat in mind, the standard approach offers a natural starting point to interpret our findings.

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, leading to a downward sloping health production frontier in the popular Grossman model (see, e.g. the analysis in Zweifel, Breyer, and Kifmann 2009). Figure 3 represents the health production frontier by the full lines. 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 those for the Mestizo in 2012 (the full gray line closest to the origin).[22] 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 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.

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.[23] Remember that this is meaningful only if the crucial (and untestable) consistency assumption holds.

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

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

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 the time-dependent control variables (altitude of district, urban status, marital status, and work status), $D2_t$ is a dummy variable indicating the follow-up wave, $P65_i$ is the treatment dummy and ε_{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 β and γ 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.[24] In particular, the individual fixed effects α_i 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:

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

This is an extended version of equation (2) that, as explained in Section 4.3, 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,

Table 5. Subjective well-being regression.

	Δ SWB		
	(1)	(2)	(3)
Treatment	−0.362	0.378	0.404
	(0.683)	(0.815)	(0.815)
Treatment × Quechua		−1.414	−1.410
		(1.603)	(1.603)
Treatment × Aymara		3.077	1.612
		(2.535)	(2.458)
Δ Expenditures	0.728	0.984*	0.976
	(0.591)	(0.594)	(0.593)
Δ Health	0.208***	0.199***	0.186***
	(0.019)	(0.019)	(0.019)
Δ Health × Aymara			0.160***
			(0.051)
Δ ADL	0.047***	0.050***	0.049***
	(0.013)	(0.013)	(0.013)
Δ Respect	0.052***	0.047***	0.044***
	(0.013)	(0.013)	(0.013)
Δ Altitude	0.803	0.849	0.734
	(1.729)	(1.673)	(1.705)
Δ Urban status	0.334	0.252	0.221
	(2.963)	(3.001)	(2.992)
Δ Marital status	1.012	1.285	1.190
	(1.212)	(1.192)	(1.199)
Δ Work status	0.014**	0.014**	0.014**
	(0.006)	(0.006)	(0.006)
Constant	0.431	−1.406**	−1.465**
	(0.524)	(0.647)	(0.647)
Constant × Quechua		3.655***	3.636***
		(1.195)	(1.194)
Constant × Aymara		5.348***	5.266***
		(1.597)	(1.559)
Observations	1775	1775	1775
R^2	0.160	0.179	0.183
R^2	0.157	0.175	0.178

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

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 are shown in Table 5. They are in line with results found in other satisfaction studies, including those with Peruvian data.[25] All four dimensions have a positive effect on SWB (yet, expenditure is significant only in the second specification). 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. From a normative point of view it is interesting to check whether the Pension 65 program has an independent effect on SWB, after controlling for the other life dimensions. This is not the case: the Pension 65 treatment has no significant effect on SWB after controlling for the other dimensions.[26]

As can be seen from column (3), only one interaction remains at the end of our iterative procedure: that is the interaction between the health index and the Aymara dummy. This significant effect suggests that the Aymara attach more importance to health than the other ethnic groups in our sample.[27] This difference in preferences is further illustrated by the results in Table 6, which shows the willingness-to-pay (WTP) for a one point increase on a non-income dimension (measured on a 100-point scale), as a percentage of expenditures.[28] 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 19.07 %. Respondents who identify as Aymara, on the other hand, are willing to reduce their expenditures much more (35.47 %) 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.

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. 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, embodying more attention for the non-material dimensions of life). 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. 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).[29] 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 recipients a shift to a healthier life style. We see indeed among the treated Aymara lower levels of nutritional deficiency, as measured by their Mini Nutritional Assessment (MNA) score (Guigoz 2006), which could reflect a shift to food with a higher nutritional quality.[30] Moreover, the improved access to health care may induce a greater awareness of health problems and an improvement of the information about health. This latter channel is proposed by Bernal, Carpio, and Klein (2017) as the main explanation for the impact of expanding health insurance coverage through the Peruvian 'Integral Health Insurance Plan'.

6. Preference differences and policy evaluation

We have now obtained all the information that is needed to implement the two welfare criteria that were introduced in Section 2. In fact, the subjective well-being criterion is easy to implement: one simply needs to check whether the subjective well-being index has increased after the treatment. The analysis in the previous section also suggests an easy test for the preference-based criterion: individual i will reach a higher indifference curve if $(\beta \triangle ln(y_i) + \gamma \triangle x_i) > 0$. Equation (4) clearly shows 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

Table 6. WTP for an increase of 1 point on the index (expressed as percentage of expenditures).

	Mestizo	Quechua	Aymara
Health	19.07	19.07	35.47
ADL	5.07	5.07	5.07
Respect	4.48	4.48	4.48

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.

Using the coefficient estimates from column (3) of Table 5, we report in Table 7 the share of the sample who are better off in 2015 compared to 2012, 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 (despite the strong treatment effect). 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 sub-group 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 7 focus on the two welfare criteria discussed in Section 2. Interestingly, we see that a majority of all groups report a higher subjective well-being, whereas only for the Aymara a majority reaches a higher indifference curve. We see for all groups a remarkable difference between both measures: the fraction that has moved to a higher indifference curve, is considerably smaller than the one reporting a higher subjective well-being. Part of the explanation may be that between 2012 and 2015 all groups have 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) and Gana et al. (2012). Another part of the explanation may be that the subjective satisfaction answers capture elements that are not taken up in our analysis of indifference curves, more specifically the social

externality that follows from caring for other family members.

This brings us to the fundamental question, that is raised by the last row of Table 7. While more respondents across all groups were lifted to a higher indifference curve in the treatment group than in the control group, we still find that between 34 and 57% of respondents were not lifted to a higher indifference curve at all. How can the introduction of a non-contributory pension, which increases opportunities for attaining a higher indifference curve, result in some beneficiaries ending up on a lower one? There are several possible explanations. First, our preference estimation with this dataset is inevitably crude, which may lead to model misspecification for certain respondents or groups. We were unable to fully account for interpersonal preference heterogeneity with the available data. Second, some respondents may have made optimization errors, meaning that while they could have reached a higher indifference curve, they instead ended up on a lower one. Lastly, and perhaps most importantly, the program may not increase opportunities for all respondents. As mentioned, our well-being measures are at the individual level, whereas decisions about how to allocate the benefit may be made at the household level. Moreover, households receiving a benefit may experience reduced transfers from other sources. A deeper analysis of these effects can be found in Olivera and Iparraguirre (2024). They estimate the intention-to-treat (ITT) effect of Pension 65 and find that the program causes a reduction of 29.8 Soles in family transfers from a baseline of 42.6 Soles at the eligibility cut-off, that is, a crowding-out effect of −70% at the eligibility threshold.[31] Regarding the effects on time allocations, no statistically significant effects are found on working hours, but they do find a substantial increase in the hours dedicated to childcare (an increase of 4.15 hours per week, compared to a baseline of 1.8 h) and a significant reduction in the time spent on leisure activities, on volunteering and on social activities. All these effects are concentrated among males. Taking all this together, it is easily seen that the program may have brought the beneficiary of the program to a lower indifference curve, while

Table 7. Percentage of respondents who are better off in 2015 compared to 2012.

Criterion	Total	Mestizo		Quechua		Aymara	
		C	T	C	T	C	T
Expenditures	54.70	55.04	56.08	49.79	64.09	32.95	47.69
Health	49.18	43.73	51.05	46.06	47.88	56.82	69.23
ADL	55.77	57.49	60.28	47.72	43.63	62.50	64.62
Respect	68.79	68.30	72.73	60.17	62.55	77.27	73.85
Subjective well-being	53.97	49.63	52.59	56.02	54.44	63.64	73.85
Preference-based	45.75	41.77	46.85	42.74	43.24	55.68	66.15

Table 8. Percentage of respondents who moved to a higher indifference curve, compared to SWB.

	Lower SWB	Higher SWB	Total
Lower indifference curve	31.94	22.31	54.25
Higher indifference curve	14.08	31.66	45.75
Total	46.03	53.97	100.00

Table 9. Difference-in-difference estimation of the effect of the Pension 65 program on the likelihood to move to a higher indifference curve (logit model).

	Higher indifference curve	
	(1)	(2)
Treatment	0.166*	0.237*
	(0.098)	(0.127)
Treatment × Aymara		0.236
		(0.359)
Treatment × Quechua		−0.196
		(0.224)
Δ Altitude	−0.248	−0.239
	(0.223)	(0.232)
Δ Urban status	−0.384	−0.343
	(0.478)	(0.485)
Δ Marital status	−0.034	−0.034
	(0.205)	(0.204)
Δ Work status	0.006***	0.006***
	(0.001)	(0.001)
Constant	−0.201***	−0.283***
	(0.076)	(0.102)
Constant × Quechua		0.043
		(0.168)
Constant × Aymara		0.567**
		(0.236)
Observations	1775	1775
R^2		
Pseudo R^2	0.017	0.024

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

still generating positive effects for other family members (on whom we have no information in the data). It may also explain why the effect on subjective satisfaction is larger for those who care about the well-being of other family members.

Table 8 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.

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 4. 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 9. We see in column (1) that, overall, the Pension 65 program treatment has a significantly positive effect on recipients. When allowing for heterogeneous time trends across the ethnic groups in the control group (in column (2)), 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 significant.

7. Conclusion

In this paper, we have shown that explicitly integrating the ethnic dimension in program evaluation is possible and may yield relevant new insights. To do that, we have first compared and discussed two criteria for policy evaluation that are sensitive to (ethnic) differences in preferences. One criterion is based on a comparison of self-reported subjective well-being scores. While this criterion is sensitive to the ordinal preferences of the concerned individuals, it is also affected by changes in aspirations and reference frames. The latter aspect is not attractive, we argue. We therefore present a second criterion that is only sensitive to ordinal preferences.

We illustrate the differences between both criteria by looking at the introduction of Pension 65, a non-contributory pension program in Peru. We find that the program had a stronger positive impact on the Aymara than on other ethnic groups. In addition, the differences in outcomes are linked to differences in preferences, suggesting 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). A striking result is that the Pension 65 program does not lift all beneficiaries to a higher indifference curve. We suggest that this finding may be explained by the crowding-out of other transfers and by the reallocation of time within the household, more specifically for men.

Our empirical analysis of Pension 65 is intended to illustrate the importance of the choice of normative criteria, rather than to provide a definitive causal analysis. In this regard, the study has several limitations. First, the dataset has small sample sizes, particularly for indigenous groups, and may suffer from selection bias, potentially leading to biased estimates. Preference heterogeneity may furthermore be related to the eligibility cut-off of the program in different ways across the ethnic groups. Finally, while the treatment and control groups are broadly comparable, treated individuals are, by the

eligibility rules of the program, poorer than those in the control group. This may challenge the plausibility of the parallel trends identification assumption and, consequently, the causal interpretation of our findings, even though we partly mitigate this concern by allowing for differential trends across ethnic groups.

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 programs. 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.[32] 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; Decancq and Nys 2021; Fleurbaey and Schokkaert 2013).

At a more methodological 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.

Notes

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. A more elaborate discussion of different approaches to measuring well-being can be found in Fleurbaey and Blanchet (2013), Decancq, Fleurbaey, and Schokkaert (2015b) and Adler and Fleurbaey (2016).
3. Our argument echoes the discussion of 'physical-condition neglect' by Sen (1985, 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'.
4. On equivalent incomes, see Decancq, Fleurbaey, and Schokkaert (2015a, 2017). Samuelson (1977), Deaton (1979) and Fleurbaey and Tadenuma (2014) provide examples of quantity metric well-being measures, Samuelson (1974) and Bosmans, Decancq, and Ooghe (2018) provide examples of money metrics. For recent surveys, see Fleurbaey and Maniquet (2011) and Fleurbaey and Blanchet (2013).
5. In principle, all individuals classified as poor by SISFOH are eligible for SIS – that is, both the extreme poor and the non-extreme poor – but it is likely to involve relatively lower participation by non-extreme poor people in SIS.
6. The SISFOH score includes information about the type of fuel used for cooking; the access to water, sewerage, electricity and telephone; the material quality 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 home overcrowding.
7. The specific problems of defining ethnicity in Peru are further discussed in Paredes (2007), Sulmont (2011), Moreno (2014) and Pasquier-Doumer and Brandon (2015).
8. The nineteenth 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.
9. 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.'
10. The 2017 census in Peru has used a self-identification approach.
11. There are exceptions to this trend. Pasquier-Doumer and 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.

12. 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.

13. 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 substantially change our findings.

14. 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 INEI to estimate official poverty rates.

15. 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.

16. Bando, Galiani, and Gertler (2020) estimate the intent-to-treat (ITT) effects of the program by comparing the outcomes between eligible and ineligible individuals in the follow-up wave, 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). We use recipient status rather than eligibility condition to identify the control and treatment groups, so that our approach estimates the average treatment effects (ATE) of the program.

17. Bando, Galiani, and Gertler (2020) find similar results when they test for the balancedness of the assignment to treatment and control groups.

18. Bernal, Olivera, and Suhrcke (2024) show the importance of this public health infrastructure for the health outcomes of the population.

19. 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.

20. Given that the two waves are approximately three years apart, this finding is indeed striking. Similarly, Van de gaer, Vandenbossche, and Figueroa (2013) find strong health effects of the Mexican cash transfer program on the health opportunities of children with an indigenous background.

21. Bando, Galiani, and Gertler (2020) 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).

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

23. See also Decancq, Fleurbaey, and Schokkaert (2015a), Decancq and Schokkaert (2016), Decancq and Neumann (2016) and Decancq and Michiels (2019) for applications to the measurement of well-being. Fujiwara and Dolan (2016) provide a critical discussion. Alternative methods to estimate preferences are based on observed behavior (revealed preferences) or contingent valuation or discrete choice experiments (stated preferences), see Da Costa et al. (2024) for a review.

24. In a context similar to ours, Beegle, Himelein, and Ravallion (2012) and Ravallion, Himelein, and Beegle (2016) use a vignette 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.

25. 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 within the household.

26. In this paper we focus on life satisfaction as an overall indicator of subjective well-being. It has been argued that it would also make sense to include individual happiness as a separate life dimension. Sen (1985, 17) writes: '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) and Adler, Dolan, and Kavetsos (2017). Unfortunately, neither a direct measure of feelings of happiness, nor a battery of binary choices is available in the ESBAM data. Moreover, the results in Table 5 suggest that the treatment has no independent effect on life satisfaction, after controlling for the other dimensions of life.

27. 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.

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

29. Already in the nineteenth century, Forbes (1870, 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.'

30. A deeper analysis of the health effects of the Pension 65 program can be found in Bernal, Olivera, and Suhrcke (2024).

31. They also estimate the effect on the probability of receiving transfers and find that the share of eligible individuals receiving family transfers from other households is reduced by 29 pp. from a baseline at the eligibility cut-off of 42 pp., implying a reduction of 69% in the probability of receiving transfers.

32. See Sollis et al. (2024) for a systematic review of participatory well-being frameworks in different cultural settings.

Acknowledgments

Disclosure statement

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. https://doi.org/10.1016/j.jebo.2017.05.006.

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

Arrow, K. J. 1951. *Social Choice and Individual Values*. Vol. 12. New York: Yale 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 (5): 216–233. https://doi.org/10.1177/0094582X14547501.

Bando, R., S. Galiani, and P. Gertler. 2020. "The Effects of Non-contributory Pensions on Material and Subjective Well Being." *Economic Development and Cultural Change* 68 (4): 1233–1255. https://doi.org/10.1086/702859.

Beegle, K., K. Himelein, and M. Ravallion. 2012. "Frame-of-Reference Bias in Subjective Welfare." *Journal of Economic Behavior and Organization* 81 (2): 556–570. https://doi.org/10.1016/j.jebo.2011.07.020.

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 (5): 2083–2110. https://doi.org/10.1257/aer.102.5.2083.

Bernal, N., M. A. Carpio, and T. Klein. 2017. "The Effects of Access to Health Insurance: Evidence from a Regression Discontinuity Design in Peru." *Journal of Public Economics* 154 (2): 122–136. https://doi.org/10.1016/j.jpubeco.2017.08.008.

Bernal, N., J. Olivera, and M. Suhrcke. 2024. "The Effects of Social Pensions on Nutrition-Related Health Outcomes of the Poor: Quasi-Experimental Evidence from Peru." *Health Economics* 33 (5): 971–991. https://doi.org/10.1002/hec.v33.5.

Bernheim, B. D. 2009. "Behavioral Welfare Economics." *Journal of the European Economic Association*7 (2–3): 267–319. https://doi.org/10.1162/JEEA.2009.7.2-3.267.

Bernheim, B. D., and A. Rangel. 2009. "Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics* 124 (1): 51–104. https://doi.org/10.1162/qjec.2009.124.1.51.

Bosmans, K., K. Decancq, and E. Ooghe. 2018. "Who is Afraid of Aggregating Money Metrics?" *Theoretical Economics* 13 (2): 467–484. https://doi.org/10.3982/TE2825.

Calestani, M. January, 2009. "An Anthropology of 'the Good Life' in the Bolivian Plateau." *Social Indicators Research* 90 (1): 141–153. https://doi.org/10.1007/s11205-008-9317-5.

Camacho, A., and E. Conover. 2011. "Manipulation of Social Program Eligibility." *American Economic Journal: Economic Policy* 3 (2): 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." Technical report, 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 (6): 1139–1144. https://doi.org/10.1093/ije/31.6.1139.

Copestake, J., M. Guillen-Royo, W.-J. Chou, T. Hinks, and J. Velazco. 2009. "The Relationship between Economic and Subjective Wellbeing Indicators in Peru." *Applied Research in Quality of Life* 4 (2): 155–177. https://doi.org/10.1007/s11482-009-9070-1.

Da Costa, S., K. Decancq, M. Fleurbaey, and E. Schokkaert. 2024. Preference Elicitation Methods and Equivalent Income: An Overview." PSE Working Paper 2024-60. https://shs.hal.science/halshs-04840652v1.

Deaton, A. 1979. "The Distance Function in Consumer Behaviour with Applications to Index Numbers and Optimal Taxation." *The Review of Economic Studies* 46 (3): 391–405. https://doi.org/10.2307/2297009.

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

Decancq, K. 2023. "Cumulative Deprivation: Identification and Aggregation." In *Research Handbook on Poverty and Inequality*, edited by U. R. Wagle, 52–67. Arizona: Edward Elgar Publishing.

Decancq, K., M. Fleurbaey, and E. Schokkaert. 2015a. "Happiness, Equivalent Incomes, and Respect for Individual Preferences." *Economica* 82 (s1): 1082–1106. https://doi.org/10.1111/ecca.2015.82.issue-s1.

Decancq, K., M. Fleurbaey, and E. Schokkaert. 2015b. "Inequality, Income, and Well-Being." In *Handbook of Income Distribution*, Volume 2 of *Handbook of Income Distribution*, edited by A. B. Atkinson and F. Bourguignon, 67–140. Amsterdam: Elsevier.

Decancq, K., M. Fleurbaey, and E. Schokkaert. 2017. "Well-Being Inequality and Preference Heterogeneity." *Economica* 84 (334): 210–238. https://doi.org/10.1111/ecca.2017.84.issue-334.

Decancq, K., and M. A. Lugo. 2013. "Weights in Multidimensional Indices of Well-Being: An Overview." *Econometric Reviews* 32 (1): 7–34. https://doi.org/10.1080/07474938.2012.690641.

Decancq, K., and A. Michiels. 2019. "Measuring Successful Aging with Respect for Preferences of Older Persons." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 74 (2): 364–372.

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*, edited by M. D. Adler and M. Fleurbaey. Oxford: Oxford University Press.

Decancq, K., and A. Nys. 2021. "Non-Parametric Wellbeing Comparisons." *European Economic Review* 133:103666. https://doi.org/10.1016/j.euroecorev.2021.103666.

Decancq, K., and E. Schokkaert. 2016. "Beyond GDP: Using Equivalent Incomes to Measure Well-Being in Europe." *Social Indicators Research* 126 (1): 21–55. https://doi.org/10.1007/s11205-015-0885-x.

Ferreira, F. H. G., and M. A. Lugo. 2013. "Multidimensional Poverty Analysis: Looking for a Middle Ground." *The World Bank Research Observer* 28 (2): 220–235. https://doi.org/10.1093/wbro/lks013.

Ferrer-i-Carbonell, A., and P. Frijters. 2004. "How Important is Methodology for the Estimates of the Determinants of Happiness?" *The Economic Journal* 114 (497): 641–659. https://doi.org/10.1111/j.1468-0297.2004.00235.x.

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

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

Fleurbaey, M., and E. Schokkaert. 2013. "Behavioral Welfare Economics and Redistribution." *American Economic Journal: Microeconomics* 5 (3): 180–205.

Fleurbaey, M., and K. Tadenuma. 2014. "Universal Social Orderings: An Integrated Theory of Policy Evaluation, Inter-Society Comparisons, and Interpersonal Comparisons." *The Review of Economic Studies* 81 (3): 1071–1101. https://doi.org/10.1093/restud/rdu006.

Forbes, D. 1870. "On the Aymara Indians of Bolivia and Peru." *The Journal of the Ethnological Society of London (1869–1870)* 2 (3): 193–305. https://doi.org/10.2307/3014456.

Fujiwara, D., and P. Dolan. 2016. "Happiness-Based Policy Analysis." In *The Oxford Handbook of Well-Being and Public Policy*, 286–320. Oxford: Oxford University Press.

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." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 68 (4): 540–552. https://doi.org/10.1093/geronb/gbs093.

Guigoz, Y. 2006. "The Mini Nutritional Assessment (MNA) Review of the Literature -- what Does it Tell Us?" *Journal of Nutrition, Health and Aging* 10 (6): 460–487.

GuillenRoyo, M. 2008. "Consumption and Subjective Well-Being: Exploring Basic Needs, Social Comparison, Social Integration and Hedonism in Peru." *Social Indicators Research* 89 (3): 545–555.

GuillenRoyo, M. 2011. "Reference Group Consumption and the Subjective Well-Being of the Poor in Peru." *Journal of Economic Psychology* 32 (2): 259–272. https://doi.org/10.1016/j.joep.2009.12.001.

Imbens, G., and D. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. London: Cambridge University Press.

Kahneman, D., and A. B. Krueger. 2006. "Developments in the Measurements of Subjective Well-Being." *Journal of Economic Perspectives* 20 (1): 3–24. https://doi.org/10.1257/089533006776526030.

Kant, S., I. Vertinsky, B. Zheng, and P. M. Smith. 2014. "Multi-Domain Subjective Wellbeing of Two Canadian First Nation Communities." *World Development* 64:140–157. https://doi.org/10.1016/j.worlddev.2014.05.023.

MIDIS. 2013. "Nota metodologica de la evaluacion de impacto del programa nacional de asistencia solidaria pension 65." Technical report, Ministerio de Desarollo e Inclusion Social.

Moreno, M. 2014. "Patrones De Autoidentificación Etnorracial De La Población Indígena En Las Encuestas De Hogares En El perú." *Debates en Sociología* 39:39–71. https://doi.org/10.18800/debatesensociologia.201401.002.

Neelsen, S., and O. O'Donnell. 2017. "Progressive Universalism? The Impact of Targeted Coverage on Health Care Access and Expenditures in Peru." *Health Economics* 26 (12): e179–e203.

Ñopo, H., J. Saavedra, and M. Torero. 2004. "Ethnicity and Earnings in Urban Peru." Technical Report 980, IZA Discussion Paper Series.

Olivera, J., and Y. Iparraguirre. 2024. "The Effects of Social Pensions on Monetary and Time Transfers Among the Poor: Evidence from Peru." Unpublished report.

Olivera, J., and I. Tournier. 2016. "Successful Ageing and Poverty: The Case of Peru." *Ageing & Society* 36 (8): 1690–1714. https://doi.org/10.1017/S0144686X15000665.

Paredes, M. 2007. "Fluid Identities: Exploring Ethnicity in Peru." Technical Report 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. https://doi.org/10.1016/j.worlddev.2015.03.001.

Ravallion, M., K. Himelein, and K. Beegle. 2016. "Can Subjective Questions on Economic Welfare Be Trusted?" *Economic Development and Cultural Change* 64 (4): 697–726. https://doi.org/10.1086/686793.

Samuelson, P. A. 1974. "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory." *Journal of Economic Literature* 12 (4): 1255–1289.

Samuelson, P. A. 1977. "Reaffirming the Existence of "reasonable" Bergson-Samuelson Social Welfare Functions." *Economica* 44 (173): 81–88. https://doi.org/10.2307/2553553.

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

Sen, A. K. 1999. *Development as Freedom*. Oxford: Oxford University Press.

Sen, A. K. 2004. "Capabilities, Lists, and Public Reason: Continuing the Conversation." *Feminist Economics* 10 (3): 77–80. https://doi.org/10.1080/1354570042000315163.

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

Sollis, K., N. Biddle, H. Maulana, M. Yap, and P. Campbell. 2024. "Measuring Wellbeing across Culture and Context -- Are We Getting it Tight? Evaluating the Variation in Wellbeing Conceptualisations throughout the World." *Social Indicators Research* 174 (1): 123–155. https://doi.org/10.1007/s11205-024-03382-z.

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 (22): 9985–9990. https://doi.org/10.1073/pnas.1003744107.

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 (1): 47–78. https://doi.org/10.1080/17442222.2011.543873.

Valderrama, J. A., and J. Olivera. 2023. "The Effects of Social Pensions on Mortality Among the Extreme Poor Elderly." Technical report, LISER WP 2023–05.

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 (2): 282–310. https://doi.org/10.1093/wber/lhs032.

Van Praag, B., and B. Baarsma. 2005. "Using Happiness Surveys to Value Intangibles: The Case of Airport Noise." *The Economic Journal* 115 (500): 224–246. https://doi.org/10.1111/j.1468-0297.2004.00967.x.

World Bank. 2015. *Indigenous Latin America in the Twenty-First Century: The First Decade*. Washington, DC: World Bank.

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

Appendix. Sample selection

We construct the final sample as follows (see Table A1). We start from an initial sample of 8089 observations and drop the following cases: 275 observations who received the transfer before the baseline survey, 188 observations with no information about their recipient status, 238 observations who are ineligible because they are already receiving another pension or are registered in any contributory pension system, 171 observations living in households classified as non-poor by SISFOH, 326 observations with no SISFOH information, and 292 observations with missing information on our variables of interest (including 24 observations regarded as outliers in expenditure).

After the mentioned selections, we have an intermediate sample of 6599 observations: 3497 respondents in 2012 and 3102 in 2015. A total of 761 observations are lost due to attrition. 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 2288 respondents with a fluid ethnicity were dropped. The final (balanced) sample consists of 1775 respondents in both waves.

Table A2 focuses on the 3497 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. 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.

Table A1. Sample selection.

	2012	2015	Total
Initial sample	4242	3847	8089
Non-eligible (treated before baseline)	143	132	275
Missing SISFOH score	97	91	188
Non-eligible (pensioner or affiliated)	124	114	238
Non-eligible (non-poor classified)	88	83	171
Missing recipient status	170	156	326
Missing data	123	169	292
Intermediate sample	3497	3102	6599
Attrition from panel	578	183	761
Fluid ethnicity	1144	1144	2288
Final sample	1775	1775	3550

Table A2. Sample selection.

	Attrition from sample (1)	Fluid identity (2)	Consistent identity (3)
Expenditures	0.092	−0.229***	0.148**
	(0.095)	(0.078)	(0.073)
Health	−0.006*	0.002	0.001
	(0.003)	(0.002)	(0.002)
ADL	−0.005**	0.003	0.001
	(0.002)	(0.002)	(0.002)
Respect	−0.003	0.003	−0.000
	(0.002)	(0.002)	(0.002)
SWB	−0.013***	−0.002	0.009***
	(0.004)	(0.003)	(0.003)
Altitude	−0.032	0.072**	−0.048
	(0.037)	(0.033)	(0.031)
Urban	0.146	0.038	−0.117
	(0.117)	(0.094)	(0.090)
Age	0.052***	−0.011	−0.021**
	(0.011)	(0.009)	(0.008)
Male	0.132	−0.519***	0.371***
	(0.103)	(0.083)	(0.077)
Marital status	−0.078	−0.034	0.081
	(0.104)	(0.084)	(0.081)
Work status	−0.004***	0.001	0.002*
	(0.001)	(0.001)	(0.001)
Quechua	−0.258**	−1.165***	1.108***
	(0.127)	(0.115)	(0.098)
Aymara	−0.420*	−2.461***	1.899***
	(0.242)	(0.328)	(0.206)
Constant	−3.857***	0.892	−0.217
	(1.053)	(0.857)	(0.816)
Observations	3497	3497	3497
Pseudo R^2	0.048	0.062	0.061

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