



Time Use and Household-Centric Measurement of Welfare in the Digital Economy

Diane Coyle

(Cambridge University)

Leonard Nakamura

(Federal Reserve Bank of Philadelphia)

Paper prepared for the 36th IARIW Virtual General Conference

August 23-27, 2021

Session 3: Time Use, Welfare and GDP I

Time: Tuesday, August 24, 2021 [14:00-16:00 CEST]

Time Use and Household-Centric Measurement of Welfare in the Digital Economy

Diane Coyle and Leonard Nakamura¹

Abstract: There is substantial interest in developing a broader understanding of progress than the standard indicator of growth in real GDP, not least because the digitalization of the economy is significantly changing production and household activity. This includes the substantial amounts of time people now spend online. This paper advocates an extended utility framework combining time allocation – over working for pay, producing at home, or leisure – with monetary measures of objective or subjective well-being during each activity. How we feel while engaged in these activities encompasses all the possibilities for well-being. Implementation would require time use statistics in addition to well-being data and direct survey evidence, such as the willingness to pay for leisure time. We advocate an experimental set of time and well-being accounts, with a particular focus on digitally-driven shifts in behavior.

JEL codes: D11, D60, I31

Keywords: Time use, digitalization, welfare, well-being

Gross Domestic Product primarily measures monetary transactions, equating inputs to outputs in nominal terms. The household inputs it counts are predominantly paid work hours, while utility is considered as household monetary expenditures on consumption goods. Although widely used as shorthand for economic progress, the limitations of GDP are well-known (Coyle 2014). An alternative perspective is Becker's (1965) full income, where all household hours are considered as inputs, and utility is a function of both time spent and monetary expenditures on goods. In this perspective, household work

¹ University of Cambridge and Federal Reserve Bank of Philadelphia, respectively. We are grateful to Penny Mealy, Charles Hulten, Solomon Tarlin, and participants in the IMF Statistical Forum, a workshop in economic welfare at the Economic Statistics Centre of Excellence, and a session at the 2019 Royal Economic Society conference for their comments. Our thanks also to anonymous ESCoE reviewers for their constructive suggestions. Coyle received funding for this work from the Economic Statistics Centre of Excellence. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

and leisure time are both inputs into the production of household utility, such that utility is a function of a household consumption technology.

There is increasing interest in a full income perspective because, thanks to digitalization, the marginal monetary price of consumption is often zero. This may be either because consumption is paid for with a barter transaction, as described by Nakamura et al (2018), or because consumption products are part of a subscription bundle. When this is true, as Goolsbee and Klenow (2006) point out, then the relevant cost that the consumer faces in choosing what and how much to consume is the shadow value of time. Competition in many digital markets is competition for consumers' attention to advertising (Anderson & Peitz 2019). Additionally, digitally-produced consumption goods, such as social media and product ratings, are increasingly produced with inputs both from firms and from households.

One way to think of the difference between the approaches is that GDP involves arm's length, monetary transactions, albeit progressively adding non-market transactions (imputed rent for owner-occupied housing, financial intermediation services indirectly measured, a range of intangible investments), whereas full income includes the shadow value of all household time, and is thus very substantially larger. Alternatively, one can consider that full income takes utility seriously. That is, it measures the full experience of an economic agent during the day, including time spent at paid work, at unpaid household work, and at leisure. Either way, it offers a more complete approach than GDP to economic welfare.

There are two perspectives on valuation of full income with very different empirical implications. One, due to Becker, is to view the shadow value of unpaid time as equal to the market wage, on the grounds that this represents the opportunity cost of leisure or of household work. This is the approach used by the UK's Office for National Statistics in its household production satellite accounts. Another is to view the shadow value of time as equal to the cost of household chores, where the price of household chores is the wage rate of household workers, the route adopted by the US Bureau of Economic Analysis in their satellite accounts. These two perspectives produce very different results,

as pointed out in Bridgman (2016), particularly as in recent decades the wage rate of household workers has fallen relative to the average wage.

Even so, these two approaches cannot, unfortunately, be seen as bounds on the true value of full income.

One reason is that the value of leisure or household work time might exceed the market wage. The monetary wage is only one of the possible gains a worker experiences from paid labor...Paid labor may have intrinsic value of various kinds, including the pleasantness of the task, the meaning of work, or on-the-job learning. Of course, the wage may also overstate the value of time if the task is unpleasant, the work is viewed as unsavory, or draining to one's human capital. Conversely, hiring a worker to perform household chores may have intrinsic costs or benefits to the employer beyond the wage paid. Households may choose to employ a household worker because they experience social benefits beyond the household chores provided, such as companionship. Can we uncover these intrinsic costs and benefits in non-paid activities?

An additional problem is that, as Hulten and Nakamura (2019) point out, the utility of a market good to the consumer is not fixed, but is affected by changes in household consumption technology. If the household consumption technology is fixed, then the purchase of a given good today has the same effect on utility as the purchase of that good in the previous period. However, improvements in the information environment change the expected utility of goods. An important aspect of information is that improvements in knowledge improve the services provided. For example, pricing doctor visits or semester hours as if they were constant quality does not take into account improvements in the science known to the doctors and professors. Similarly, restaurant ratings and reviews may improve a consumer's ability to better match their tastes to dining options. Enhanced information raises utility without changing the good provided or its supply cost. Indeed, any increase in the precision of a consumer's actionable information raises expected utility. Furthermore, in this era of network economies, network externalities change the user value of social media, ecommerce platforms, and so on, over time. Just as a telephone has greater value when more households possess telephones, so the expected utility of any

network good rises as it is more widely adopted, although the good does not change.

Given these considerations, there are several possible approaches to estimating economy-wide full income. One is to delve more deeply into self-reports of well-being, developing measures of how much economic agents enjoy alternative activities. Time use studies with subjective modules are available across a variety of countries and time periods. Ultimately, these might either bound the value of activities, or lead to direct monetary evaluations of subjective states. A second approach is to look to self-reports of choices of different possible activities, where economists are increasingly looking to surveys to understand time allocation, especially if monetary compensation for behavior changes is possible to ensure incentive compatibility. And a third approach is to use parametric models with econometric measurement. Ultimately, these methods need to be combined to obtain some estimates of full income.

The base methodology for the measurement of 'real' GDP is to first create nominal GDP accounting for all monetary transactions (plus some imputations) in the economy, and then to deflate it using period-to-period changes in prices of well-defined products. The theoretical rationale is that the deflation methodology approximates the use of an expenditure function (measuring the cost in today's prices of purchasing last period's utility). Thus deflated GDP is a constant-utility construct.

To consider how to develop measures of nominal and real (constant-utility) full income, instead, we begin with subjective utility (Kahnemann 1999) as the sum of (time-separable) utility over time, $\sum_t U(t)$. This accounts for all of an economic agent's time. For leisure, this utility is then thought of as involving market goods, x , purchased with wage, w , and other income, that are combined with non-market goods produced in the household, y , and leisure time. During marketed work and household work, this utility captures the intrinsic enjoyment (or its dislike) associated with the activities of production, including the meaning attached to the activity (such as self-expression). An additional output of activity is learning (and its inverse, human capital obsolescence).

Full income can be considered to be the sum of money-metric utility over time. This includes both the monetary and non-monetary transactions already included over time, and it requires inferences about the shadow value of time measured in the same money-metric. In Becker's (1965) simplest model, utility is consumption of household commodities, which are created using market goods combined with time needed for preparation and consumption. This time is evaluated by the market wage. A key question is whether the market wage is the correct shadow value of time. Becker's simplifications ignore the portion of utility experienced directly during household and wage labor production. Moreover, by identifying the shadow value of time with the average wage, it ignores the complications due to the constraint of a standard work-week causing the marginal value of labor to diverge from its average value. Still, this framework is more likely to capture a full picture of economic welfare in a digital economy with its zero price goods (Brynjolfsson et al 2019), increased involvement of the household in the economy outside of wage labor (Coyle 2019), and rapid advances in the application of information (Hulten and Nakamura 2019).

I Direct measurement of well-being in time spent

In the standard approach, utility maximization ideally should combine all these choice margins: the consumer's choices of market hours, home production hours, leisure, and commodities, subject to the time identity and the usual budget constraint. However, this ambitious extension with utility depending also on the extended time required for production (market and home) and consumption (with leisure and consumption as joint activities), may not have any of the general equilibrium characteristics with which we are familiar (Steedman, 2001).

An alternative approach is to directly consider well-being in the dimension of time use. There is a large and growing literature on the measurement of the well-being derived from different activities. Our basic proposition is that utility over time is equivalent to well-being. How we feel while working for pay, producing at home, or at leisure encompasses all our possibilities for well-being. For a real-terms measure, we can ask, as we do with dollars, how much time it would require to achieve the same utility attained in the previous period.

Indeed, time spent offers a potentially more equitable way of valuing nonmarket goods. Asking people how much they would be willing to pay for something is always skewed by how much income they have (just as markets will overly represent rich people's preferences). There is new interest in measuring household income growth more democratically (Aitken & Weale, 2018). But since time is the great leveler, asking people how much time they would be willing to spend could provide more equitable valuations. A democratic measure of full income might be built around time units rather than monetary units.

To consider how to implement the well-being in time spent framework, we start by thinking about a world in which the shadow value of time is equal to the market wage, and hours are fully adjustable. Suppose that digitalization makes leisure time more valuable in well-being terms. It is possible that market hours worked could fall (if the income effect outweighs the substitution effect) and wages could rise as less labor were offered, raising its marginal productivity. But the reverse could also happen.

One way to capture these effects might be to take time use data and ask participants for evaluations of their well-being during different activities, as is done in time use surveys. Alternatively, we could ask participants to report their monetary valuations of different activities, in effect the consumer surplus that they experience receiving. These types of studies have been used in cost-benefit analysis of government-provided free goods, so there is a well-developed literature (for example, Viscusi, 2018, Small, 2012). Moreover, a series of papers have argued that recent increases in the availability of data on time use provide a robust path forward for the measurement of household production (Aguiar et al, 2012, Aguiar et al 2018, and Aguiar and Hurst, 2007, 2016).

We would expect such self-evaluations of either kind to be changing as digitalization is causing relative price changes in terms of time as well as money, and could be expected to lead not only to shifts in expenditure and consumption patterns but also to changes at the work/leisure/home production margins. There have been obvious changes in the time/money combinations involved in the purchase and consumption of some services. For example, time saved walking to the bank and standing in line may be used in watching a video online. More time

may be spent taking and looking at smartphone photos, at negligible cost, or creating and uploading songs written for pleasure at home, and less on dining out. Perhaps there is less home cooking and more ordering food through delivery apps.

A challenging feature is that substitutions of this kind may be hard to pin down through time use studies. Mobile apps often work in the background, giving us reminders, instructions, messages, and information while we are doing other things. In particular, the availability of many possible actions via a smartphone makes it particularly useful in periods of downtime such as waiting or queuing. This may turn periods that would otherwise be ones of boredom and discontent into active leisure, in effect creating new leisure time out of thin air. Self-reports are one way to explore these dimensions. In principle, time use surveys can capture the primary and alternative activities people are engaged in at a given time, although this is clearly somewhat harder than ascertaining whether somebody is ironing and watching TV at the same time.

Absent new time use data, it is impossible to be sure about systematic aggregate changes, but since the launch of the first smartphone in 2007 use of the mobile Internet has become an ever-present activity in many people's lives. This has enabled the rapid growth of new services, from social media to digital apps and platforms, as well as new channels of distribution and access. An estimated 80 percent of all people over the age of 14 in developed economies were connected to the mobile Internet, mostly by smartphones, by December 2016 (Evans, 2017). The available statistics indicate substantial growth in the volumes of data transmitted over mobile and fixed networks during the past decade, with a doubling in mobile data usage in 37 countries in the year to December 2017 alone (<http://www.oecd.org/sti/broadband/broadband-statistics/>).

As Hulten and Nakamura (2017) show, the pervasive Internet has an important impact on the nature of the transmission of information. The transmission of actionable information increases the value of any economic activity, including consumption. The fact that information can be transmitted online at close to zero marginal cost changes behavior and prices in ways that make the measurement of economic growth difficult (Hulten and Nakamura, 2019; Coyle, 2019). Hulten and Nakamura (2019) and Nakamura et al. (2018) provide evidence that production

measures of output growth may be an order of magnitude smaller than welfare measures for specific innovations.

II Issues in measuring well-being

The contrast between asking a general question (as in Juster 1985) and a specific retrospective time period question (as in Krueger 2006) is related to Kahneman's (1999) distinction between "objective" and "subjective" utility. For objective utility, we want to know how an experience feels in real time. It is evident that our recollection of the past may differ from our moment-to-moment feelings. Gershuny (2013) and Krueger (2009a) consider self-reports on the enjoyment experienced during different activities, such as at work, driving in traffic, or at leisure out of the home. Gershuny deploys mean activity enjoyment scales, while Krueger uses unhappiness indexes, measured as the proportion of time during the event when negative feelings are rated as strongest. Both are based on diary self-reports, as opposed to the expensive studies where individuals are asked to report in real time in response to random signals. However, Krueger et al. (2009a) present evidence that, on average, remembered feelings, as measured in their survey, are reflective of moment-to-moment feelings, as detected in surveys conducted with special devices for recording feelings at specific points in time. This is an ongoing area of study, and it is possible that progress can be made since the use of mobile devices for reporting may enable low-cost extensions of these surveys. Empirical evidence of this kind may also bring us closer to understanding how experienced utility at work and the wage rate are related to the shadow value of leisure.

Extensive studies by behavioral economists and psychologists on decision-making suggest that we often do not maximize utility when making decisions but rather follow rules of thumb. How does this affect the welfare value of consumption revealed by purchasing decisions? Benjamin et al. (2012) asked individuals to choose between alternative scenarios, such as having a lower rent (20 percent of income) and a longer commute (45 minutes) or a higher rent (40 percent of income) and a shorter commute (10 minutes). Moreover, they asked the same individuals whether they believed this choice would lead to higher life satisfaction, greater happiness with life as a whole, or greater felt happiness (subjective well-

being). They find that there are systematic differences between the choices people say they would make and what would maximize these various definitions of happiness. They also find that higher life satisfaction is most aligned with choice, while subjective well-being is less so. And while some of the difference between choice and what would actually maximize happiness is explicable as a problem of self-control, self-control issues appear to explain only a minority of the differences.

Benjamin et al. (2014) looked at the same question — how does choice relate to what would make us happiest on various definitions — in a real world context by examining the residency training choices of medical graduates and asking how their choices would affect happiness and life satisfaction during the residency as well as how they would affect attainment of the best possible life. They find that none of these three measures well explained people's choices. In particular, the residency prestige and the desirability of the choice for the individual's partner had considerable additional weight. Thus, professional and social concerns — above and beyond what seems most desirable for the individual — are important in understanding choice.

Another issue is whether a single dimensional measure such as happiness is the appropriate way to measure episodic utility. Krueger et al. (2009a) use five dimensions of feeling and combine them to distill an overall measure of time spent in unpleasantness; a time period is unpleasant when the strongest feeling experienced is negative (stressed, in pain, or sad, as opposed to happy or interested). This allows for the fact that, for example, an episode of work may contain more elements of pain or stress than, say, watching television. Can these multidimensional feelings in fact be placed in a single monetary metric as Krueger et al. suggest?

For that matter, are scaled self-reports associated with specific activities, whether single dimensional or multidimensional, in turn relatable to scaled self-reports of overall happiness, as in the Cantril scale? To the extent that the Cantril scale can be related to log measures of income, it may be possible to apply meaningful monetary values to specific activities. In turn, we might be able to associate these feelings with actual expenditures. That is, when someone pays to attend a rock

concert or for a meal, do their feelings line up with their expenditures? Or are the feelings we experience and report partly mediated by the size of our outlays? On the other hand, Kahneman and Deaton (2010) provide evidence that Cantril scale reports and emotional well-being scales are less well correlated for higher incomes, which would limit the value of this strategy.

One possibility is to use stated preferences to predict out-of-sample behavioral consequences, as suggested by Bernheim et al. (2013). They advocate using econometric techniques to measure the extent to which revealed preferences are predicted by stated preferences. This may help us to understand the extent to which a given monetary measure reflects actual choices. On the other hand, we might believe that feelings — as revealed by feelings in time use studies — are more definitive of welfare. Either way, to the extent we can reconcile the results of different methods, the more confidence we can have in them.

There are several additional challenges in implementing the measurement at an aggregate level of well-being across activities. For the links between activities in the different conventional economic categories and average well-being in undertaking them are not straightforward. There are several complexities.

First, surveys capture subjective reports of how respondents feel — emotions such as happiness or anxiety. This surely cannot be taken directly as a measurement of utility, as subjective contemporaneous feelings need to be supplemented by purpose for longer-term goals such as health, education, child rearing, or entrepreneurship. One possibility would be to consider such goals as an investment in individual capital, or capability and social goods, contributing to others, or to the next generation.

Moreover, subjective reports will differ across individuals. How an individual scores feelings will contain random elements, possibly both person-specific and time- or context-specific. One way to deal with this is to treat these reports as a dependent variable with proxies for true utility on the right hand side, as in Blanchflower and Oswald (2004).

Another caveat is that work can be enjoyable or not, depending on the job, and, even when intrinsic job satisfaction is low, there are benefits from the social

attachments and status that come with paid employment. There is evidence that the nonmonetary aspects of work are significant, and people seek intrinsic meaning in their paid work (Cassar and Meier, 2018). What's more, the (dis-)utility of work appears to be changing over time as the character of work changes, and there are also substantial variations between groups (Kaplan and Schulhofer-Wohl, 2018; see also Jahoda, 1981, on the "latent" value of work).

Some home production activities are similarly enjoyable and blend with consumption/leisure, while others are clearly "chores" (Gershuny and Fisher, 2014). Paid labor or household production can both be directly very enjoyable or welfare enhancing. In effect, paid labor or household production can tip over into leisure. Leisure can also be productive. While we are at leisure, we can come up with good ideas or upload content that others may enjoy and learn from. Sichel and von Hippel (2018) argue that household research and development is substantial relative to private research and development.

Finally well-being, on the standard Cantril scale, is measured relative to the 'best possible' life. The best possible life changes over time due to economic innovation. That is, novel economic possibilities, such as greater longevity, deeper scientific understanding, tastier food, and more captivating entertainment, may change the definition of the best possible life. This will affect the measurement of well-being over time.

Despite these complexities, to a first approximation we might think that time reductions (holding output constant) in paid labor and home production — that is, in what we call 'work' — are an improvement in welfare. Conversely, increases in time working (either in home production or for pay) given constant output are, in principle, welfare worsening. This is on the assumption that, either in home production or at work, the object being produced is the major purpose of that time. Thus, while one may enjoy washing dishes or writing essays, one would prefer to do these tasks in a shorter period of time rather than stretching out these episodes.

For leisure, the presumption is the opposite: To a first approximation, the more time allocated to it, the better. It is likely that for many activities there are

diminishing returns. How much time one spends at a given activity depends on how rapidly the returns diminish. On the other hand, in general, more time spent at leisure suggests more enjoyment per unit of time for that activity. This is the hypothesis that underlies the Goolsbee and Klenow (2006) analysis of the Internet. Of course, unemployment is a bad (forced) “leisure” in that it restricts our ability to obtain the highly productive goods of the marketplace, which may force us back toward the less productive branches of home production. And this overall low level of productivity likely further lowers the enjoyment of leisure time, as we are denied the goods we are accustomed to consuming at leisure.

III Monetary measures of well-being

It is important to note that it is when *holding income constant* that reductions in time spent in either paid labor or household production are leisure- and welfare-enhancing. (While unemployment may result in more time spent at leisure, it is not welfare-enhancing.) However, holding income constant implies that *a monetary measure is required*.

The utility measures that we propose are, at least potentially, provided with a quantitative metric because of their connection to the consumption and production of goods. How far we can proceed down this road is above all an empirical question. Although Nordhaus (2009) argues that the data cannot exist to proceed along this path, Krueger et al. (2009b) deliver a spirited rejoinder. While Krueger et al. (2009a) opt not to proceed to the natural conclusion of a monetary welfare measure, they argue that it is feasible.ⁱ Instead, in their attempt to integrate aggregate time use figures with well-being results in a “National Time Accounting,” they calculate a national well-being index that tracks changes over time resulting from changing time use patterns among different population groups. National Time Accounting of this kind produces a measure supplementing conventional GDP figures but is not a monetary metric. For the reason explained previously, we seek to draw the lines connecting potential monetary valuations and personal consumption expenditures, on the one hand, and measures of consumer self-reports on feelings and happiness, on the other.

Is it therefore possible to assign shadow prices to these feelings and thus to different uses of time? There are several options. Essentially these correspond to

the broader debate about the relationship between stated preference, stated feelings, and revealed preference. Economists are inclined to place more weight on revealed preference measures, but a good deal of policy-oriented welfare analysis rests as well upon stated preferences as providing valuable additional evidence. An excellent example of this can be seen in Small's (2012) discussion of the valuation of travel time as a crucial input into any cost-benefit analysis of transportation policy. He discusses travelers' stated valuation of travel time costs and compares it to their preferences as revealed, for example, by econometric analyses of commuting time-rental tradeoffs. He points out that the evidence for the welfare impact of in-vehicle amenities is thin. Amenity questions will become even more salient as we realize the possibility of partially or totally self-driving cars. Reported measures of happiness or other feelings while driving may help bridge this gap.

One route would be to ask survey participants directly about their shadow value of time, just as Brynjolfsson et al. (2019) ask about the monetary value of different digital consumption/leisure activities (see also Coyle et al, in progress). Such studies introduce monetary scales of utility in the evaluation of goods, asking how much subjects would be willing to pay for a given amenity (such as social media) or how much they would be willing to accept to do without the amenity.

However, although such contingent valuation studies are widely used in environmental economics, the more usual approaches to self-reports of utility in the context of the well-being literature are based on arbitrary scales. The best-known of these are the happiness studies, where subjects are asked to report, for example, in terms of the Cantril ladder, how they rate their lives currently on a scale of 0 to 10 with respect to the best possible life they could be leading. While this scale is both arbitrary and context-specific, Deaton (2008) and Stevenson and Wolfers (2008) show that responses across countries are on average well approximated by a linear regression on log income per capita. So self-reports of utility appear to be, at least in cross section, relatable to a cardinal, monetary measure of utility.ⁱⁱ

It is true that the studies in Kahneman (1999) show that such self-reports are affected by many factors other than measured real income. But can self-reports be

placed on a monetary scale? The self-reported monetary values reported by Brynjolfsson et al. (2019), based on willingness to pay/willingness to accept methodologies and thus related to compensating and equivalent variation, suggest that there may be some way to do so. Surveys could ask: What would people be willing to pay for an extra day's vacation, provided their workloads were reduced? What would they have to be paid to work an extra day, assuming their workloads were not reduced? What would they pay for someone else to perform a household chore or at what pay would individuals work an additional hour at their current jobs or at some benchmark alternative? The answers to such questions could then be related their wage rates and the measured, experienced utility of labor.

In an alternative approach, not reliant on contingent valuation methods, Bridgman (2016) uses estimates of the replacement cost of household activities to derive a first version of a household production account. Since the average wage rate for household employees across types of work does not vary very much, we can easily approximate the value of household production if we assume that hired labor is a reasonably good substitute for home production. This approach assumes that the shadow price of time for highly paid workers can be equated to the wage rate of household employees. But if highly paid workers are, say, deeply concerned about their children's education and/or enjoy their interactions with their children, then the shadow price of their time may be substantially higher. The former implies greater household production, but as investment, while the latter adds to consumption. Diewert et al. (2018) show how to estimate the shadow price of household production using the wage and the wage rate of employees as well as the case when neither wage rate is applicable.

Alpman et al. (2018) take yet another approach, using experienced well-being and time use surveys combined with money measures to estimate directly the monetary value of nonmarket activities. In essence, they scale money expenditures with estimates of experienced well-being (along the lines of Krueger et al., 2009a) within a representative agent framework to estimate total income for a range of countries. Their approach is somewhat ad hoc, not being derived from a full-blown theory of individual welfare. Yet they are able to link time use and well-being ratings to different activities to estimate the relative "well-being"

valuations of nonmarket activities and then multiply these by total consumption expenditure to derive a monetary measure of welfare. This is an important first step in using experienced well-being surveys to estimate the value of the shadow price of time.

IV Issues in measuring the shadow value of time

As discussed, a shadow price of time is needed to construct a monetary measure of full income. If we were to ask workers how much they would require to work an extra hour at a ‘neutral’ job — one that, say, requires some concentration but is not stressful — the difference between the pay they would demand for this compared with their current job could be a metric of the utility cost (or benefit) of their work. Pay at the ‘neutral’ job should reflect the true marginal value of leisure. This would be analogous to the standard use of hedonic wage regressions in order to isolate the marginal benefit or disbenefit of certain job characteristics as compared with average wages.

Maestas et al. (2018) ask workers about their preferences for working conditions, such as flexibility in hours, vacation time, and meaningfulness of the work, and how much they would be willing to accept in pay reductions to change them. This enables them to discuss the extent to which working conditions exacerbate wage inequality. The answers will likely also reflect the shadow value of time. Mas and Pallais (2017) ask similar questions in the context of call centers, where they can also measure the revealed preferences of the workers.

An additional question raised by Cassar and Meier (2018) is whether the experienced utility measures that we use are adequate for capturing nonmonetary incentives that may affect the shadow value of time. In particular, they argue that the meaningfulness of labor, particularly as captured in the mission or purpose of the work, has an important impact on the pay workers are willing to accept for a given task. They point to a variety of empirical evidence in the human resource management literature that bears on this question. The willingness of better-paid workers to work longer hours may be due to the meaningfulness of their labor as well as subsidiary factors such as autonomy, competences, and social relatedness. It may also be the case that child rearing, which is experienced as often being

unpleasant, may have a “mission” attached to it that more than makes up for short-term disutility.

To the extent that happiness can be translated into a monetary metric, changes in happiness can be interpreted as equivalent to changes in real income. For example, Blanchflower and Oswald (2004) use 1972–1998 data from the U.S. to calculate that for males it would take some \$60,000 (1990 dollars) to compensate them for being unemployed; that is, an unemployed male would be as happy as an employed male with similar income if given an additional \$60,000. This probably cannot be interpreted as measuring the direct utility of employment in addition to the wage, as the status of being unemployed is different from being employed in “neutral utility” work (Frey, 2008, pp. 45–53). However, similar techniques might be used to interpret the sort of data presented in Krueger et al. (2009a) or Juster (1985), which could be more closely tied to a concept of “neutral” utility if these data could be related to the overall happiness of individuals. This is a conceivable complement to surveys that ask about the monetary value of the utility of work.

V Intrinsic well-being at work

The assignment of monetary values to the well-being associated with different uses of time will need to overcome some complications. Earlier, we noted that a key complexity concerns the enjoyment people can derive from work. This question has nagged at economists since the studies that underlie Juster et al. (1981) first revealed how many workers value their work. This is consistent with the emphasis in the positive psychology literature on “flow,” or satisfying absorption in a meaningful activity (Nakamura and Csikszentmihalyi, 2002). As a first approximation, economists including Becker (1965) have argued that the hourly wage is the opportunity cost of an hour of leisure. This assumes that the work itself is neither pleasant nor unpleasant. But some people have jobs they enjoy quite a lot, while others report that they find their work relatively unpleasant. Rothwell and Crabtree (2019) provide survey evidence that job satisfactions beyond the wage are important to workers and correlate with reported well-being. The value of leisure depends then on both the wage the worker receives and how much intrinsic utility they get from that job. This may

depend on whether labor conditions in the economy as a whole have been stable; it is possible that the average utility of labor as experienced has changed, perhaps due to a change in the bargaining power of workers (as evidenced by the declining labor share; see, for example, Bental and Demougin, 2010). The composition of labor has also changed, which could also change the aggregate utility of experienced labor, as argued by Kaplan and Schulhofer-Wohl (2018).

Other changes may be occurring. If the population is experiencing greater distress, as suggested by Case and Deaton (2017) and Deaton (2018), it is of value to explore how paid work might be contributing to this. Equally, the utility people receive from different types of nonmarket production may vary; for example, Lerner and Tirole (2003) suggest that developers of open source software gain three types of utility: enjoyment from the activity, peer esteem, and future rewards in terms of pay and promotion in their career. Juster et al. (1981) and Juster and Stafford (1991) have argued that a more complete welfare accounting might include the underlying utility experience at both paid work and household production.

The evidence to date on the enjoyment of work is mixed. In the 1975 and 1981 surveys presented in Juster (1985), employed Americans were asked to record their level of enjoyment of 20 activities on a scale as bounded by 10 (enjoy a great deal) and 0 (don't enjoy at all), with 5 representing an activity to which the respondent was indifferent. As shown in Table 1, the activity "job" was given a mean score of 8.02 in 1975 and 7.79 in 1981. The next *more* enjoyable category was "going on trips, outings" with the ratings 8.24 in 1975 and 8.17 in 1981, while the next *less* enjoyable category was "home entertainment" with the ratings 7.76 (1975) and 7.54 (1981). By contrast, cooking was rated 6.17 (1975) and 6.13 (1981), and television 5.93 (1975) and 6.00 (1981). While one might object that the high rating for paid work and the low ranking for household work reflects the wage received for the former, sorting the jobs by occupation results in "virtually no association between the process benefits from work and the intrinsic characteristics of the job as reflected by its occupational status," (Juster, 1985, p. 341).

On the other hand, in the Princeton Affect and Time Survey Krueger et al. (2009a) asked respondents to record their happiness on a scale of 0 to 6, where 0 meant not experiencing the feeling at all and 6 meant the feeling was very strong. These data are presented in Table 2. (They were asked the same question about feeling stress, sadness, interest, and pain.) In sharp contrast to the Juster (1985) studies, work was among the least enjoyable activities, well below all leisure time activities and quite similar to or worse than household production activities.

Whether this reflects the substantial differences in methodology between the two surveys or differences in the experience of work is unclear. For example, the Juster (1985) survey asked about general attitudes toward activities, while the Krueger (2009a) survey asked about specific episodes in the previous day. However, if it is the case that employment has become substantially more, or less, pleasant, then this would have first-order effects on monetary estimates of utility.

VI The way forward

The agenda of measuring broader economic welfare in terms of a money metric of the well-being afforded by different allocations of must address several open questions, discussed here, in order to progress. This is in addition to the underlying issue of the need for more detailed and regular time use data, including digital activities. These open questions concern respectively the measurement of well-being and of the shadow value of time in monetary terms.

We have set out a series of questions about the linkages among measures of utility, consumption expenditures, and time allocation to work and leisure, and about the measurement of the shadow value of time. These research questions derive from the earlier seminal work on time use by Becker, Gershuny, Kahneman, Juster, Krueger, and many others. This distinguished tradition is given new urgency not only by the current public debate about the inadequacy of conventional real GDP as a measure of economic welfare or progress but also by the evident significant changes in time use in both consumption and production processes due to digital innovation.

These open questions relate first to how our different activities relate to our overall evaluations of our current well-being and secondly from these evaluations to money measures of economic activity. Some of these questions may be answered by econometric studies, while others may be answered through survey methods. There is therefore a rich research agenda concerning the meaning of self-reports on different methodologies, the utility derived from different activities at leisure and at work, the best approach to applying a money metric, and the potential need for more than one dimension to measure economic welfare. One thing that is already clear is the need for updated and more frequent time use surveys.

There is, finally, a wider question about time use as the foundation for an economic welfare metric. This is whether it points to consideration of a capabilities-based rather than utility-based approach to social welfare. Capabilities (Sen, 1982, 1993) refer to what people can *do* rather than what they can spend, although their possibilities for spending will constrain what they can do. In social welfare terms, it is not really the economic outputs themselves that concern us, and yet in GDP economists have constructed a measure based on expenditure and output, imperfectly adjusted through deflation to link to underlying utility. Our leisure and work activities are enhanced by our capabilities. But many of our capabilities are unexercised in a particular period of time or may never be exercised. To the extent that capabilities make us able to contemplate larger or more productive outcomes, they are reflected in our work activities. Social interaction, purpose, and capability may be alternative dimensions to how we are feeling. For example, having a “serious” conversation with someone may not be joy-filled, but it may be more valuable than joy. Having a purpose may require actions that provoke deep anxiety. However, this goes beyond our scope here.

The effort to come up with an additional measure of economic well-being is unlikely to have as well-defined or uncontroversial a quantification as current measures of GDP until this research agenda is much further advanced. Agreement on measurement is more likely to come about if we examine economic well-being through multiple lenses and work toward an understanding about the most convincing ways to measure it. How might this quantification be established as a

long-term means of evaluating a national economy's contribution to the welfare of its residents? Macroeconomists and policymakers currently rely upon GDP or some of its components to answer this question. If there is an increasing gap between the answer supplied by measures of GDP and measures based on welfare, then it may be that a measure of welfare should become part of the system of national accounts. Establishing this additional accounting may be crucial if economists are to be able to discuss economic policy issues meaningfully, in a context in which there is growing public questioning of whether real GDP growth is an adequate measure of broad economic progress. However, this task will require a sustained dialogue between government statisticians and the economics profession at large.

The 2020-21 pandemic raises further questions, both for the measurement of GDP and for the measurement of well-being. Health outcomes, in a world in which some countries' health care expenditures can exceed ten percent of GDP, are an increasingly important part of measured real growth. At the same time, the relationship between such real expenditures and either direct health outcomes or well-being are not closely connected. Health outcomes, as the pandemic has shown, are the outcome of health shocks and of prior health expenditures and accumulated human capital. As a consequence, well-being may be worsened by health shocks, regardless of the efficacy of the health care system. Such shocks would, in principle, be registered as a decline in measures of full income. These costs include excess deaths, millions of COVID and long COVID patients, and the psychological and educational tolls of isolation, fear, and disruption, as well as direct economic costs.

In addition, the unprecedented speed with which the pharmaceutical firms and governments were able to develop, trial, approve, and manufacture vaccines that are highly protective against the new coronavirus and its variants is a credit to advances in the world economy and evidence of its puissance. This appears to be allowing much of the world to return to more normal levels of activity, a feat that could be valued as worth trillions of dollars.

How should we incorporate these events into both the time series of GDP and full income? In current SNA practice, we generally do not see either the full,

extraordinary costs of the pandemic or the extraordinary economic benefits from the innovative ideas, development and distribution of vaccines. The metrics include only the expenditures on the development of the vaccine made by pharmaceutical companies and governments, and then the costs of producing and distributing the vaccine, largely borne by governments.

The rise in global temperatures and other ecological impacts associated with climate change raise similar questions about how to develop a measure of well-being that incorporates all relevant considerations. For example, in the formulation developed by Partha Dasgupta in *The Dasgupta Review: The Economics of Biodiversity* (2021), Nature impacts the economy as a set of assets or resources and also as a direct influence on the environment in which consumption and other economic activities take place.

Both types of challenge illustrate the growing wedge between standard national accounts measures and well-being, and further underline the case for new approaches to measurement.

References

Aguiar, Mark, Mark Bilal, Kerwin Charles, and Erik Hurst, 2018. "Leisure Luxuries and the Labor Supply of Young Men," Princeton Working Paper, February. <https://scholar.princeton.edu/maguiar/publications/leisure-luxuries-and-labor-supply-young-men>

Aguiar, Mark, and Erik Hurst, 2007. "Measuring Trends in Leisure: The Allocation of Time Over Five Decades," *Quarterly Journal of Economics* 122 (3), August, 969–1006.

Aguiar, Mark, and Erik Hurst, 2016. "The Macroeconomics of Time Allocation," in John Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis, 2012. "Recent Developments in the Economics of Time Use," *Annual Review of Economics* 4, 373–397.

Aitken, Andrew, and Martin Weale, 2018. "A Democratic Measure of Household Income Growth: Theory and Application to the United Kingdom," Economic Statistics Centre of Excellence Discussion Paper 2018-02.

<https://www.escoe.ac.uk/wp-content/uploads/2017/02/ESCoE-DP-2018-02.pdf>

Alpman, Anil, Fabrice Murtin, and Carlotta Balestra, 2018. "Unveiling the Monetary Value of Non-Market Activities Using Experienced Well-Being and Time Use Surveys," paper prepared for the 35th IARIW General conference. <http://www.iariw.org/copenhagen/alpman.pdf>

Anderson Simon P and Martin Peitz (2019), 'Ad Clutter, Time Use and Media Diversity,' University of Mannheim Discussion Paper 140, December 2019.

Becker, Gary S., 1965. "A Theory of the Allocation of Time," *Economic Journal* 75, 493–517.

Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Alex 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.

Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Alex Rees-Jones, 2014. "Can Marginal Rates of Substitution Be Inferred from Happiness Data? Evidence from Residency Choices," *American Economic Review* 104 (11), 3498–3528.

Bental, Benjamin, and Dominique Demougin, 2010. "Declining Labor Shares and Bargaining Power," *Journal of Macroeconomics* 32 (1), 443–456.

Bernheim, B. Douglas, Daniel Bjorkegren, Jeffrey Naecker, and Antonio Rangel, 2013. "Non-Choice Evaluations Predict Behavioral Responses to Changes in Economic Conditions," NBER Working Paper 19269.

Blanchflower, David G., and Andre J. Oswald, 2004. "Well-Being over Time in Britain and the USA," *Journal of Public Economics* 88 (7–8), 1359–1386.

Bridgman, Benjamin, 2016. "Accounting for Household Production in the National Accounts: An Update, 1965–2014," *Survey of Current Business* 96 (2), 1–5.

- Brynjolfsson, Erik, Avinash Collis and Felix Eggers, 2019. Using massive online choice experiments to measure changes in well-being, *Proceedings of the National Academy of Sciences*, 116 (15) 7250-7255; DOI: 10.1073/pnas.1815663116
- Case, Anne, and Angus Deaton, 2017. "Mortality and Morbidity in the 20th Century," *Brookings Papers on Economic Activity* Spring, 397–443.
- Cassar, Lea, and Stephan Meier, 2018. "Nonmonetary Incentives and the Implications of Work as a Source of Meaning," *Journal of Economic Perspectives* 32 (3): 215–238.
- Coyle, Diane, 2014. *GDP: A Brief but Affectionate History*, Princeton NJ: Princeton University Press.
- Coyle, D., 2019, Do-it-yourself Digital: the Production Boundary, the Productivity Puzzle and Economic Welfare. *Economica*, 86: 750-774. doi:[10.1111/ecca.12289](https://doi.org/10.1111/ecca.12289)
- Coyle, Diane, David Nguyen and Joel Rogers De Waal, "Free digital goods and economic welfare," in progress.
- Dasgupta, Partha, *The Economics of Biodiversity: The Dasgupta Review*. <https://www.gov.uk/government/collections/the-economics-of-biodiversity-the-dasgupta-review>, February 2021.
- Deaton, Angus, 2008. "Income, Health, and Well-Being Around the World: Evidence from the Gallup World Poll," *Journal of Economic Perspectives* 22:2, 53–72.
- Deaton, Angus, 2018. "What Do Self-Reports of Wellbeing Say About Life-Cycle Theory and Policy," NBER Working Paper 24369, March.
- Diewert, W. Erwin, Kevin J. Fox, and Paul Schreyer, 2018. "The Allocation and Valuation of Time," paper prepared for the 35th IARIW General Conference. <http://www.iariw.org/copenhagen/fox.pdf>

Evans, Benedict, 2017. "Mobile Is Eating the World," Andreessen Horowitz presentation, <https://www.slideshare.net/a16z/mobile-is-eating-the-world-20162017/3-3Were-halfway-to-connecting-everyone55bn>

Frey, Bruno S., 2008. *Happiness: A Revolution in Economics*, Cambridge, MA: MIT Press.

Gershuny, Jonathan, 2000. *Changing Times: Work and Leisure in the Postindustrial Society*, Oxford University Press.

Gershuny, Jonathan, 2013. "National Utility: Measuring the Enjoyment of Activities," *European Sociological Review* 29 (5), 996–1009.

Gershuny, Jonathan, 2018. "Gender Symmetry, Gender Convergence and Historical Work-Time Invariance in 24 Countries," University of Oxford Centre for Time Use Research Working Paper, February. https://www.timeuse.org/sites/default/files/2018-02/CTUR%20WP%202%202018_1.pdf

Gershuny, Jonathan, and Kimberly Fisher, 2014. "Post-Industrious Society: Why Worktime Will Not Disappear for Our Grandchildren," Oxford University Sociology Working Papers, 2014-03, April. <https://www.sociology.ox.ac.uk/materials/papers/wp20143.pdf>

Goolsbee, Austan, and Peter J. Klenow, 2006. "Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet," *American Economic Review* 96 (2), 108–113.

Hulten, Charles, and Leonard Nakamura, 2017. "Accounting for Growth in the Age of the Internet: The Importance of Output-Saving Technological Change," NBER Working Paper 23315, April.

Hulten, Charles, and Leonard Nakamura, 2019. "Expanded GDP for Welfare Measurement in the 21st Century," working paper. <https://www.nber.org/chapters/c13886.pdf>

Jahoda, Marie, 1981. "Work, Employment, and Unemployment: Values, Theories, and Approaches in Social Research," *American Psychologist* 36 (2), 184–191.

Juster, F. Thomas, 1985. "Preferences for Work and Leisure," in F. Thomas Juster and Frank P. Stafford, eds., *Time, Goods, and Well-being*, Ann Arbor, MI: Survey Research Center, University of Michigan, 333–351.

Juster, F. Thomas, Paul N. Courant, and Greg K. Dow, 1981. "A Theoretical Framework for the Measurement of Well-Being," *Review of Income and Wealth* 27 (1), 1–31.

Juster, F. Thomas, and Frank P. Stafford, 1991. "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement," *Journal of Economic Literature* 29 (2), 471–522.

Kahneman, Daniel, 1999. "Objective Happiness," in Daniel Kahneman, Ed Diener, and Norbert Schwartz, eds., *Well-Being: The Foundations of Hedonic Psychology*, New York: Russell Sage, 3-25.

Kahneman, Daniel, and Angus Deaton, 2010. "High Income Improves Evaluation of Life But Not Emotional Well-being," *Proceedings of the National Academy of Sciences* 107(38), September 10, 16489-93.

Kahneman, Daniel, Ed Diener, and Norbert Schwartz, eds., 1999. *Well-Being: The Foundations of Hedonic Psychology*, New York: Russell Sage, 3-25.

Kaplan, Greg, and Sam Schulhofer-Wohl, 2018. "The Changing (Dis-)utility of Work." *Journal of Economic Perspectives* 32 (3), 239–258.

Krueger, Alan B., Daniel Kahneman, David Schkade, Norbert Schwartz, and Arthur A. Stone, 2009a. "National Time Accounting: The Currency of Life," in Alan B. Krueger, ed., *Measuring the Subjective Well-Being of Nations*, University of Chicago Press: Chicago, 9–86.

Krueger, Alan B., Daniel Kahneman, David Schkade, Norbert Schwartz, and Arthur A. Stone, 2009b. "Rejoinder," in Alan B. Krueger, ed., *Measuring the Subjective Well-Being of Nations*, Chicago: University of Chicago Press, 243–251.

Lerner, Josh, and Jean Tirole, 2003. "Some Simple Economics of Open Source," *Journal of Industrial Economics* 50 (2), 197–234.

Maestas, Nicole, Kathleen J. Mullen, David Powell, Till von Wachter, and Jeffrey B. Wenger, 2018. "The Value of Working Conditions in the United States and Implications for the Structure of Wages," NBER Working paper 25204, October.

Mas, Alexandre, and Amanda Pallais, 2017. "Valuing Alternative Work Arrangements" *American Economic Review* 107 (12), 3722–3759.

Nakamura, Jeanne, and Mihalyi Csikszentmihalyi, 2002. "The Concept of Flow," in C. R. Snyder and S. J. Lopez, eds., *Handbook of Positive Psychology*. New York: Oxford University Press, 89–105.

Rothwell, Jonathan and Steve Crabtree, Not Just a Job: New Evidence on the Quality of Work in the United States, 2019,
<https://www.gallup.com/education/267590/great-jobs-lumina-gates-omidyar-gallup-quality-report-2019.aspx>

Sen, Amartya, 1982. *Choice, Welfare and Measurement*, Oxford: Blackwell.

Sen, Amartya, 1993. "Capability and Well-Being," in Martha C. Nussbaum and Amartya Sen, eds., *The Quality of Life*. Oxford: Clarendon Press.

Sichel, Daniel, and Eric von Hippel, 2018. "Household Innovation, R&D, and New Measures of Intangible Capital," working paper, November.

Small, Kenneth A., 2012. "Valuation of Travel Time," *Economics of Transportation* 1 (1–2), 2–14.

Steedman, Ian, 2001. *Consumption Takes Time: Implications for Economic Theory*, Routledge.

Stevenson, Betsey, and Justin Wolfers, 2008. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox," *Brookings Papers on Economic Activity*, Spring, 1–87.

Table 1. Basic Process Benefits (Juster, 1985)

Activity	1975 Score	1981 Score
Talking with children	9.16	8.98
Care of children	8.87	8.74
Trips with children	8.87	8.72
Games with children	8.62	8.24
Talking with friends	8.38	8.27
Going on trips, outings	8.24	8.17
Job	8.02	7.79
Home entertainment	7.76	7.54
Reading books, magazines	7.60	7.49
Going to church	7.23	7.28
Reading newspapers	7.17	7.10
Making things for house	6.78	6.47
Playing sports	6.76	6.23
Going to movies, plays	6.65	6.38
Gardening	6.55	6.27
Cooking	6.17	6.13
Television	5.93	6.00
Other shopping	5.69	5.30
Housing repairs and alterations	5.11	4.94
Work, school organizations	5.00	5.13
Grocery shopping	4.57	4.55
Cleaning house	4.22	4.18
Sleeping	NA	7.54
Eating at home	NA	7.46
Personal care	NA	7.38
Eating out	NA	7.33
Taking naps	NA	5.20
Caring for other children	NA	4.53

Activity	Happy (Raw)	Happy (after removing individual fixed effects)
Sports/exercise	5.08	4.89
Religious	4.97	4.81
Socializing	4.74	4.68
Child care	4.63	4.59
Eating/drinking	4.57	4.49
Telephone	4.47	4.50
Relaxing/leisure	4.34	4.35
Lawn/garden	4.23	4.21
Volunteer	4.22	4.28
Shopping	4.11	4.15
Travel	4.05	4.06
Food prep/cleanup	4.02	4.02
Personal care	4.02	4.07
TV	3.91	4.00
Working	3.80	3.83
Medical care	3.64	3.76
Education	3.62	3.55
Housework	3.55	3.56
Adult care	3.54	3.50
Household management	3.50	3.70
All	4.13	4.13

ⁱ "In principle it is possible to estimate the monetary price that people are willing to pay on the margin ... For example, the way workers trade off pay for a more or less pleasant job Alternatively, the amount that people are willing to spend on various types of vacations can be related to the flow of utility they receive Although it

is possible ... to put a dollar value on W in this framework, we shy away from this step ..." (Krueger et al., 2009a, p. 15). See also Gershuny (2000), chapter 8.

ⁱⁱ Because the frame for the Cantril ladder is "the best possible life," the definition of the best possible life evolves over time with new discoveries. It is less evident that these happiness measures correlate with measured real GDP over time. Benjamin et al. (2012) ask students whether they would choose to have been born about when they were (1990) or in 1950; 87 percent would choose their actual date, which contrasts with the Cantril ladder results indicating that well-being has remained flat over time.