
Towards genuine progress on the Genuine Progress Indicator

Michael Harris

Agricultural and Resource Economics,
University of Sydney, Australia
E-mail: m.harris@usyd.edu.au

Abstract: The construction of the Genuine Progress Indicator (GPI) is viewed through the lens of economic methodology. The criterion of ‘consistency with a definition (of income)’ as a basis for providing microfoundations for the GPI is contrasted with that of being derived explicitly from a formal model, with the deficiencies of the former highlighted. The importance of formal modelling and, through it, establishing causal relationships is vital to the extent that the GPI is used to draw inferences about thresholds (and more generally about the benefits vs. costs of economic growth) – which is to say, the GPI is used as a signalling device, rather than simply as a tracking device. Issues of substitutability between types of capital, thermodynamic constraints and the importance of prices are also addressed. In particular, the importance of substitutability in consumption, as opposed to simply in production, is emphasised.

Keywords: GPI; genuine progress indicator; Hicksian income; Fisherian income; sustainability; economic growth; national accounting; methodology.

Reference to this paper should be made as follows: Harris, M. (2008) ‘Towards genuine progress on the Genuine Progress Indicator?’, *Int. J. Environment, Workplace and Employment*, Vol. 4, No. 1, pp.82–94.

Biographical notes: Michael Harris is a Senior Lecturer in Resource Economics and the Head of the Agricultural and Resource Economics group in the Faculty of Agriculture, Food and Natural Resources at the University of Sydney, Australia. He has previously worked at La Trobe University and the University of Melbourne. His key research interest is in natural resource accounting (also known as ‘environmental accounting’ or ‘green accounting’) and the construction and interpretation of measures of sustainability. He is also keenly interested in the compatibility of sustainability measures and decision support systems such as benefit-cost analysis. He has recently worked on a project with researchers at CSIRO and the Beijer Institute of Ecological Economics (Sweden) on incorporating ecological resilience into wealth-based sustainability measures. In 2008, he joined the editorial board of the journal *Environment and Development Economics*.

1 Introduction

I thank Phil Lawn for inviting me to reply to his response to my article ‘On income, sustainability and the ‘microfoundations’ of the Genuine Progress Indicator’. To discuss his response paper requires going into some length, because what is under discussion is

not just the mechanics of the construction of one particular measure and its consequent interpretation. The core of our disagreement goes beyond method and into methodology – the business of how one does economics in the first place. If one takes seriously Jacob Viner's comment that economics is (defined as) what economists do, then clearly economists do a variety of things in a variety of ways. And if one takes seriously the comparatively postmodern position of McCloskey and others (see, e.g. Hausman, 1989; Randall, 1993), economics does not proceed unfalteringly by means of theoretical improvements and rigorous hypothesis testing, but by rhetoric and persuasion (backed, at least in part, by theoretical reasoning and analysis of available evidence).

Since the Genuine Progress Indicator (GPI) and its sibling measures have come under criticism for lacking 'theoretical microfoundations' (Neumayer, 1999; Neumayer, 2000), a charge which Lawn has attempted to meet head on (Lawn, 2003), it is worth examining why this charge has been made, and discussing what the role of theory 'should' be in the construction of sustainability indicators. In my previous paper in this exchange, I noted that economic theory was not a machine for generating 'truth', but rather a means for formalising enquiry (Harris, 2007, p.128). Any piece of economic theory starts from a few axiomatic principles and then, put simply, messes about in an interesting direction to see what happens. (Put slightly more formally, 'messaging about' means 'building (or writing down) a model'.) Theorising imposes logic and internal consistency on the process of enquiry: claims about the results of the theorising arise directly *from* the theorising. Any *asserted* claims are simply *ad hoc*.

Of course, *ad hoc* claims are not totally forbidden in the academic world (we all have hunches), but they are what they are: assertions made by an analyst, backed by an implicit appeal to authority.

Part of what I hope to achieve in this exchange is a more nuanced understanding of the criticism of the GPI on theoretical grounds. What theoretical underpinnings should a sustainability measure have? (Or an economic welfare measure, for that matter?) And why do 'mainstream' (for want of a better designation, although perhaps 'neoclassical' is appropriate) environmental economists remain sceptical of measures like the GPI? Methodological stubbornness? Ideological intransigence? Blind-spots resulting from having been indoctrinated into the Tribe of the Econ?¹

I hope to make clear that the reservations that various 'mainstream' economists have about the construction and calculation of the GPI are primarily methodological and analytical. That is, they arise from some basic disagreements about how economics should be done. In the spirit of postmodern relativism, where economics is viewed as a discourse, or series of discourses, then alternative discourses within economics are equally admissible. Combining this with Viner's statement that economics is what economists do, then the GPI, being something *done* by economists, must then be the result of a valid exercise in *doing* economics.

I am not that much of a postmodernist. I will devote much of this discussion talking about issues in 'doing economics', hopefully to make clear why my interpretation of how economics should be done leaves me dissatisfied with the GPI as a useful construction of applied economics.

2 Lawn's position summarised

To commence critiquing Lawn's defence of the GPI, I need to try and fairly articulate the position I intend critiquing. I interpret Lawn's overall position as being comprised of the following statements:

- 1 The GPI is an experiential measure of current economic welfare, broadly defined. It is not designed/constructed to be an explicit measure of economic sustainability.
- 2 There is an important distinction between the optimal macroeconomic scale of an economy, and its maximum sustainable scale.
- 3 The GPI is a useful indicator in regards to optimal macroeconomic scale (balancing benefits and costs of economic growth).
- 4 Since optimal macroeconomic scale will be exceeded before the maximum sustainable scale is reached, downturns in the GPI can provide an 'early warning system' for unsustainability (notwithstanding point 1).
- 5 Ecological economists do not believe (for various reasons) that produced and natural capitals are substitutes that can be meaningfully conflated within a single measure. (Net National Product (NNP) does combine the two, while the GPI as per Lawn does not.)
- 6 Maintenance of natural capital is inherently limited by the law of entropy, which means that natural capital *must* run down over time. (Hence 'sustainable growth' is an oxymoron, and the aim should be a steady-state economy.)
- 7 The GPI is adequately conceptually underpinned by a concept of income associated with Irving Fisher.
- 8 The NNP measure is associated with Hicksian income, but misleading as a sustainability indicator for reasons to do with conceptual problems with 'weak sustainability'. (A series of examples is used to show how NNP can fail to give appropriate indication of aggregate sustainability or lack thereof.)

I will be more interested in directly addressing the latter four points, given that the first four claims are dependent on the validity of the GPI, while the latter four are arguments given to *support* its validity, particularly relative to other measures in the sustainability literature.

3 Models, definitions and identities

The debate in progress is, in large part, to do with economic theory and microfoundations.² To talk sensibly about theoretical microfoundations requires us to clarify distinctions between *definitions*, *identities* and *models* in economics. (If it is not obvious why we should sort out the distinctions, it will become clearer when we revert to discussing the construction and interpretation of the GPI.) We proceed by example.

For useful examples of definitions in economics, we can look at definitions directly related to the issues under discussion here. Income is a concept subject to (various) definitions. Income as defined by Fisher is the immediate flow of psychic well-being, while Hicksian income is (broadly) defined as the standard of living that can be conceived without lowering the achievable standard of living in the future. This

oversimplified discussion of definitions shows how thorny matters can get when different people legitimately define key concepts differently. Not only did Fisher and Hicks arrive at quite different definitions of income, but as I previously noted (Harris, 2007), Hicks himself explored a number of variations to his fundamental definition.

These variations in the *specifics* of the definitions come prior to Hicks' attempts at actually *operationalising* the concept of income in a way consistent with the underlying definition (his 'central criterion') – Hicks' distinction between income *ex ante* and *ex post*. (I remarked previously on how disconsolate Hicks sounded regarding the prospects of successfully representing his central criterion empirically.) We could add further to definitional complexity by referring to Friedman's distinctions between permanent and transitory income, and broadening our definitions of income to account for uncertainty. Clearly, multiple definitions of a key central concept in economics are available.

But wait, there's more. What is referred to in the national accounts as 'national income' is defined differently again, in ways consistent with the construction of identities, so it is to identities we now turn.

Accounting identities are central to the construction of national accounts, and two of the most important identities are the 'income/production' GDP identity and the balance of payments identity. The first of these is what all economics undergraduates meet in the guise of the circular flow of income, where the value of national production necessarily equals the income earned by all factors of production.

The second identity says that the balance of payments must balance (meaning the current and capital accounts must offset each other, adjusting for government currency purchases). This identity links to the previous one in that the current account is directly tied to the (im)balance of national savings and investment. Note that these identities are linked together based on appropriate *definitions* of the key variables. It is possible to use just these linked identities to gain some insights into the effects of (for example) protectionist policies.

Two comments about identities are important to make at this point. First, identities and equalities are not the same thing. Somewhat confusingly, Lawn repeatedly refers to 'national income identities' when the expressions he refers to are equalities, *not* identities (the GPI, as constructed, is not, ever, an accounting identity³). More discussion on this below. The second comment is that while one can use identities as indicated above to provide some insights on, for example, how the impact of trade policy interventions will be constrained by identity relationships within the national accounts (so that, for example, one cannot consider the balance on current account without reference to the capital account, and this in turn requires us to consider domestic saving and investment behaviour), there are limits to how much analysis can be undertaken by solely referring to identity relationships. Without *models* of domestic saving and investment behaviour, we can do little more with respect to the balance of payments than paint in broad brush-strokes. More precise analysis must be model-driven. So now we turn to models.

In economics, a model is a conceptual laboratory for conducting thought experiments. Typically built upon a combination of axiomatic foundations and specific assumptions, the model can be tweaked to see what it looks like (for example, what characterises its optimum or its equilibrium conditions), and how it responds to changes in key variables or parameters. A conventionally constructed model in economics involves some combination of objective function and technical constraints, although within that broad specification a wide variety of specifics are admissible.

One linked pair of reasons models matter in discussions of sustainability is that, firstly, they indicate something about how the analyst thinks the world works (in the specification of the model) and, secondly, they indicate what the analyst thinks is important (by their choice of objective function, constraints, sustainability criterion and so on). Other analysts can then take issue, present an alternative model and explore its implications.

The second reason models matter in sustainability analysis is why they matter generally; that while the analyst gets to set up the model, they don't get to second-guess the answer. It is a useful discipline on what economists do that they are prepared to be surprised by the results of their modelling, not only empirically but theoretically.

The problem in the context of the GPI and its sibling measures, by the conventional standards of 'doing economics', is that no process of formal modelling has been entered into. The GPI, according to Lawn (2003), has been provided with theoretical microfoundations by being rendered *compatible with a definition*, rather than being *derived from a well-articulated model*. This is not normally what would be regarded as adequate 'microfoundations' in mainstream economics. There are many widely varying indicators that could conceivably be constructed consistent with Fisherian principles, so it is hard to say why Lawn (or anyone else) would regard the GPI as the 'correct' one. The individual elements (and their weights) used in the construction of the GPI are *ad hoc*, necessarily, chosen on subjective grounds without reference to an underlying model that might provide a guide.⁴

It might be counterargued that, even if the GPI is not constructed from a formal theoretical model, neither is GDP. However, GDP is carefully constructed around national accounting identities which provide a formal framework for making decisions about what counts and why, and how the aggregate is to be interpreted.

This not the case with the GPI, notwithstanding repeated references in Lawn (2003) to 'national accounting identities' that are in fact no such thing. For an accounting identity to hold, component elements must satisfy consistent definitions designed to ensure that (in the case of GDP) the measured product on the left-hand side is *definitionally equal* to the sum of the expenditures on the right-hand side. The circular flow identity only holds subject to appropriate definitions of variables that make up either side of the identity (based on underlying accounting principles). In the formal 'green accounting' literature, writers (see, e.g. Hamilton, 1994) are careful to distinguish between the NNP concept (where the underlying identity can be maintained because depreciation can be deducted from both output [the value of physical production] and income [profit] hence representing an equal deduction on both sides of the circular flow identity), and the Measure of Economic Welfare concept (involving adjustments to consumption that render it *not* an accounting identity).

The GPI, like the measure of economic welfare (MEW), is similarly not based on an accounting identity, but on a constructed equality. The aggregate value on the left-hand side is simply the sum of whatever has been chosen to go on the right-hand side. There is no identity relationship that constrains how the sum is put together.

4 Interpretation of the GPI (again)

As discussed in my previous paper, the question emerges as to how to interpret changes in the GPI over time. The idea of an optimal macroeconomic scale suggests that (if the optimal scale is indicated by a peaking of the GPI) the GPI is capturing changes in the

net benefits of growth. (Recall that by construction, the peaking of GPI is expected to precede the breaching of maximum sustainable scale.) To be plausibly linked to economic theory, linkages must be spelled out between the *fact* of observed economic growth (measured by GDP) and its social consequences.

This is pretty much incontrovertible in the case of (say) industrial pollution that results from the production of consumer goods. It is less obvious that marital breakdown is *caused* by economic growth, or even that every instance of divorce reduces social welfare. Appeals to the idea of Fisherian income resolve nothing, because, as noted previously, all that provides us with is consistency with a definition, but no rationale for particular choices made in the construction of the index.

This brings us back to the issue of the arbitrary choice of components in the GPI, and their subjective weighting in the final index. Clive Hamilton attempts to meet these charges head on (Hamilton, 2007, p.315) by arguing that all components are chosen in response to identifiable and well-understood deficiencies of GDP, which are weighted using market prices or dollar-valued costs (such as health costs resulting from ozone depletion).

This is highly debatable. Even if it is widely agreed that increasing income inequality is undesirable, there is no obvious or theoretically compelling weighting system using (say) Gini coefficients for adjusting growth in consumption downwards. Secondly, as with items like marital breakdown or alcohol consumption, it is not even theoretically apparent that all marriage breakdowns are social welfare reducing, let alone that increasing rates of marital breakdown are caused by faster economic growth (such that the rate of marriage breakdown would decline, with slower economic growth). Similarly, how should one deal with alcohol consumption? One might wish to adjust aggregate consumption based on tangible costs, such as estimated productivity losses due to alcoholism, and external costs (reflected in the legal and health systems) from alcohol abuse and violence associated with drinking. However, this is not how Lawn and his co-authors have dealt with consumption of alcohol.

We should also note here that Clive Hamilton is unambiguously supportive of the Hicksian criterion of sustainability as an underpinning for the GPI.

“It seems agreed that the GPI does not function as an indicator of current welfare and of sustainable income, but as an indicator of sustainable welfare. In other words, it measures what we might call ‘Hicksian welfare’, the maximum amount of welfare that a nation can enjoy over some time period and still be as well off at the end of the period as at the beginning.” (p.313)

This leads Hamilton to take a more conventional approach to accounting for changes in capital stocks than does Lawn. Where Lawn tries to include services from capitals in a way that reflects his fundamental belief in non-substitutability, Hamilton simply argues that aggregating changes in capitals implies nothing at all about their substitutability.

“Adding the value of haircuts to the value of oranges in calculating GDP does not make them substitutes for each other.” (p.311)

This claim strikes me as simply incorrect, or at the very least a misunderstanding of why we would add such things together in a measure of welfare. If I am forced to endure a temporary loss in haircutting services, I may be able to tolerate my increased shagginess a little better by consuming more citrus products. Silly though this sounds, Lawn rightly emphasises that consumption of *stuff* is a means to an end (subjective well-being), and both oranges and haircuts are ways by which we can increase our well-being. In the final analysis it is substitution in the utility function that matters, a point we shall return to.

Before moving on, it is worth making a distinction about the purpose of alternative indicators. Some indicators are primarily used as **tracking devices**, while others are used as **signalling devices**. Conventional GDP is primarily used as a tracking device, monitoring a nation's (market sector) economic activity over time. Its signalling properties are limited to short- and medium-term issues such as employment and inflation, where more formal models of aggregate supply are incorporated to solidify the analysis.

Sustainability measures are, by definition, signalling devices. Lawn stresses that the GPI is not first and foremost a sustainability indicator, but that it does serve to indicate threshold breach, which comes prior (and leads) to the exceeding of the maximum sustainable scale. Drawing such inferences from the time series of the GPI, as Lawn does, requires the spelling out of causal linkages between changes in conventional economic growth and the elements that they affect that lead to the breaching of optimal and maximum economic scales. This is the reason I have stressed the role of theory, and the explicit identification of causal linkages in my comments.

However, Lawn's commentary slips between emphasis on the purely empirical role of the GPI as a tracking device measuring changes in true current economic welfare, and its interpretive role as a signalling device, in which downturns are assumed to *mean* something about the path the economy is on. If the GPI is purely about tracking a time series, then Lawn can reasonably dismiss my concerns about causal linkages. If he wants to draw inferences from downturns in the series about thresholds being exceeded, he cannot.

5 Capital substitutability and sustainability

While Lawn stresses that the GPI is not designed to be a sustainability measure proper, a fundamental part of our discussion concerns what constitutes a meaningful measure of sustainability, and the relationship of such a measure to underlying concepts of income. Since mainstream economists lean naturally to capital-theoretic measures in the weak sustainability tradition (such as Genuine Savings), a key area for debate is the extent to which different kinds of capitals can be substituted for in the production process. Lawn argues – largely by anecdote – that natural and produced capitals are complements rather than substitutes. Taken to a global extreme, this implies something very strong about economic growth since industrialisation, and particularly about the role (or lack of) of technical change, namely, that *all* economic growth must have resulted from an enormous mobilisation of material inputs (including natural resources and labour). Technical change does not result in substitution possibilities between natural and produced capital; rather, it results in greater efficiency in digging up (or chopping down or harvesting) natural resources and using them to produce output.

There is considerable scope to disagree with this view. It is just as easy to find anecdotes involving natural and produced capital as substitutes as to find examples where they are complements. The more important matter is whether at some aggregated level, it is meaningful to think of natural and produced capital as being substitutable to some degree in terms of generating a social standard of living. Simple aggregate production functions posit an aggregated stock of manufactured capital that can be substituted for some aggregate stock of natural capital in production of a flow of output. In reality, things are of course complex; technical change typically results in changes in the

composition of a combination of produced and natural capitals that can result in certain outputs. The output itself may not be a physical good. (If I wish to have a meeting with distant colleagues, a telephone conference call or video link will involve a very different kind of technology and have very different resource use implications than if we all travel to meet in person at a central location.)

It should also be noted that models that represent ‘a stock of natural capital that is depleted over time’, as many simplified neoclassical growth models do, are oversimplifying in another way. Much human activity over time has involved not a loss of natural capital over time (as, say, with the depletion of a pool of oil), but instead a transformation of natural capital into some other form of natural capital (such as clearing land for agriculture). This has implications for the value of goods and services generated by this transformed natural capital as compared to its untransformed state, as well as for the inputs needed to keep it productive in one state compared to the other. But it is not as simple as saying that ‘natural capital has declined’ – in the case of farmland, an ecosystem has been transformed. The values generated by it have changed.

Further note that the standard of living we are generating is increasingly ‘disembodied’, and produced by comparatively disembodied means. One key growing element of the total capital stock is *human capital*, which has a production function largely unrelated to the incremental depletion of natural capital, and yet is key to generation of ongoing standards of living (assuming that technical change does in fact matter). Moreover, a lot of what we consume is becoming comprised of non-rival goods and services (note the ‘bits and bytes vs. atoms and molecules’ distinction talked about in the context of the so-called weightless economy – see Coyle, 1997), that can be consumed at the margin without using much in the way of physical resources. (Think of music sold in digital form via iTunes.) Sustaining a standard of living is not the same as sustaining a flow of a portfolio of *physical* outputs.

To repeat what I said previously, ‘the important substitution possibilities are in consumption: can individuals maintain non-declining standards of living even as some parts of the portfolio of goods, services and amenities (that we consume) are declining towards zero’ (Harris, 2007, p.121). To put it in another way, what is the degree of substitution between capitals, not in terms of producing some level of output, but in terms of generating some amount of *value*? As mentioned before, it is the substitution possibilities in the utility function (rather than the production function) that matter. Since vital ecosystem services – like clean air – are hard to substitute out of in a consumption bundle, this still leaves all of us in some agreement that some particular capitals, and the services they provide, will require special attention in any accounting or measurement system we construct.

In Lawn’s discussion of capitals and substitutability, there are two ‘elephants in the room’ (one more hidden than the other) – thermodynamics and prices. So to these, we now turn.

6 The two elephants in Lawn’s room

Economists are used to working with the idea that life is ‘subject to constraints’. Without constraints, there would be no scarcity and no need for the discipline of economics. At the most aggregated level possible (that of the universe), the ultimate constraints are provided by the laws of thermodynamics. Lawn and other ecological economists make

much of the limitations imposed upon us by thermodynamic laws, particularly the fact of entropy. It is entropy that ecological economists invoke to argue for a transition from a growth-oriented economy to a steady-state economy.

It is worth recapping, in a very non-technical form, what the first two laws of thermodynamics imply. Economists would best understand the first law as a conservation law, like an accounting identity in an economic model. Matter and energy can be converted, in principle, one for the other, but not reduced in total. (Matter–energy simply exists – it is neither created nor destroyed.)

The second law, concerning entropy, is recognisable to economists as a statement of opportunity cost, or the ‘no free lunch’ principle. Any conversion or transformation of matter–energy involves some loss along the way (entropy increases), hence Lawn’s position that continued conversion of natural capital to produced capital cannot be relied upon as a recipe for continued economic growth.

If this were all there was, Lawn’s position would seem to be unassailable. However, as already noted, maintaining a standard of living is not identical to maintaining a flow of physical outputs. Moreover, we note that the biosphere is not a closed thermodynamic system. If we think of the economy’s ‘job’ as to generate a (continuing) standard of living, then essentially over time we, as a society, are engaged in the activity of converting sunshine into a standard of living. Over time, we will (necessarily) have to move from a reliance on tapping long-stored sunshine (in the form of fossil fuels), to converting the *current* flow of sunshine into a standard of living.

Even Lawn’s steady-state economy relies on the continuing injection of sunshine into the earth’s thermodynamic system, otherwise entropy would make maintaining even the steady state impossible.⁵ Hence, the *existence* of entropy does not, in and of itself, provide an open-and-shut case for moving to a steady-state economy. Continuing with some form of economic growth will depend on the efficiency with which we can convert sunshine (the input into the global thermodynamic system) into a standard of living, noting that not only the means of conversion will change over time, but also the composition of our final standard of living will also change over time. This is *not* the same as converting sunshine into a sustained amount of output, implying a non-declining throughput.

The second ‘elephant in the room’ in Lawn’s discussion is the concept of relative prices, central to much of economics. The role of prices is hidden in much of what Lawn presents. In several examples where NNP and GPI are calculated and compared, various kinds of outputs and capitals are aggregated, implying (constant) relative prices, but the process by which prices are identified or derived is not made clear. (There are instances where Lawn refers to ‘units’, which are obviously multiples of quantity and price, but each of these is buried within the aggregated ‘unit’.) It is not clear what role the prices are playing. We might (for example), simply think of prices as being meaningful as indicators of relative scarcity (see Lawn, 2000), but as containing no information relevant to sustainability, because they do not reflect constraints imposed by thermodynamics.

If we think about *market* prices, then this remark has some force. Market prices may not only not reflect sustainability constraints (including but not limited to thermodynamic ones), but they may not even capture relative resource scarcity well if some form of market failure prevails.

However in terms of constructing an ideal index, market prices are not the relevant prices to consider. Not being constrained by the national accountant’s need to maintain national accounting identities (for which market prices will matter), the practical

economist can regard market prices as empirical approximations at best. The theorist, concerned with deriving principles for construction of a 'correct' sustainability indicator, can work entirely in terms of conceptual ('shadow' or 'accounting') prices that emerge from well-specified constrained optimisation problems. In other words, prices emerge from models.

In this context, shadow prices can capture sustainability criteria, if sustainability constraints are properly articulated in the set-up of the optimisation problem. It seems unclear to me that we know the full import of thermodynamic constraints on the earth's biosphere, while it is subject to a continuing inflow of sunshine, but if such constraints *could* be properly articulated, then shadow prices that reflected long-run entropy constraints could, in principle, be derived and used in the construction of economic sustainability indices.

Lawn's comparisons of NNP and GPI are done in the absence of formal modelling, with 'consistency with a definition' substituted in place of 'being derived from a formal model.' NNP is presented as being consistent with (a general form of) Hicksian income, while the GPI is argued to be consistent with Fisherian income. Deriving both measures from well-specified models grounded in theory could well yield different results from those that Lawn presents, not least because the prices used would not be the same. As an example, examine Lawn's fishing economy in which returns from fishing are reinvested into produced capital (production of more boats, meaning increased fishing effort). Lawn's argument is that even if the flow of consumption (periodic harvest of fish) is non-declining over the time-frame under analysis, and the stock of boats is increasing, so that measured NNP should be rising, we could be witnessing a fundamentally unsustainable economy. The increased stock of produced capital results in increased fishing effort and increased pressure on the natural resource, which (if this pattern continues) will collapse at some point in the future.

One *could* quibble with Lawn about this stylised example. It is designed as a cautionary tale of how NNP might mislead when produced and natural capital stocks are complements rather than substitutes, and hence, why one should be careful of assuming that Hartwick Rules or maintain-non-negative-Genuine-Savings prescriptions will yield sustainable outcomes. One could quibble that the unfortunate results are the consequence of it being artificially designed at a low level of aggregation, with only one form of produced capital interacting with only one kind of natural capital so that the outcome is more or less inevitable. Adding many more kinds of individual capitals into each category might well change the outcome dramatically. Lawn might respond that investing other kinds of capitals in other sectors simply shifts the problem elsewhere, putting pressure on other parts of the resource base. This returns us to the unresolved question, discussed earlier, of the substitutability of different kinds of capitals at high levels of aggregation.

However, the point here is not to quibble with the example, but to take it seriously and examine it more formally than Lawn does. Lawn's 'naïve NNP' measure simply adds current consumption to net (rather than gross) investment. His point, in short, is that net product will not indicate sustainable prospects for the economy if the depletion of the fish stock, while being 'compensated for' (using unspecified prices) by an increased build-up of 'boat capital', is leading to a situation where the fish stock will eventually be entirely depleted, leaving the boat capital valueless and the economy in collapse.

There is in fact a growing theoretical literature dealing with how to account for major, potentially catastrophic, events, including collapse of resource stocks or regime changes in ecosystems; see, for example, Mäler, Xepapadeas and de Zeeuw (2003), and Nævdal (2006). Ecologists have written for some time about ‘ecological resilience’ (described intuitively as the ability of a system to withstand shocks; see, for example, Walker et al., 2004), and economists have begun bringing resilience into economics as a variable that can be quantified, priced and incorporated into measures of overall wealth (Mäler, 2008). In Lawn’s example, the running down of the fish stock does not simply reduce wealth because there are fewer fish; it further reduces wealth because resilience is now lower (conversely, there is greater risk of catastrophic collapse).

The point of this is to note that Lawn’s NNP, which he links to Hicksian income and is at pains to demonstrate weaknesses in, is something of a straw man. It is possible (and important!) to extend the capital-theoretic approach to sustainability measurement to account for the potential for major events induced by human action. Lawn’s example does not identify the prices at which natural capital (rents from fishing) are converted into produced capital, but a fully developed resilience analysis would show overall wealth declining over time if it is the case that the collapse in the fish stock is inevitable.

To conclude this section, Lawn makes much of thermodynamics and very little of prices. But in this context, one cannot talk about the former without dealing with its impact on the latter. If Lawn can articulate *in detail* (quantitatively) how thermodynamic constraints will affect production and consumption possibilities in the future, then the next obvious step would be to write down the model that reflected these constraints explicitly. The shadow prices that emerge from that model would be appropriate to use in indicator construction.

However, if Lawn (and/or others) *cannot* articulate thermodynamic constraints explicitly and quantitatively, then we have nothing to go on. It only remains for those working in the (‘weak sustainability’) capital-theoretic tradition to continue in the direction they largely have been: modifying the initial simple measures they have been developing, by figuring out ways to deal with major events, uncertainties and tipping points and regime shifts. Approaches that have been directed to this end include the specification of safe minimum standards, the precautionary principle and resilience analysis.

7 Conclusions

It is still unclear, based on comparing Lawn’s position to that of Clive Hamilton, whether the GPI is thought to be most consistent with income definitions associated with Fisher (as per Lawn) or Hicks (as per Hamilton). To the extent that Lawn, in his response to me, viewed his task to be to show how Hicksian income had weaknesses in the measurement of sustainability, perhaps he should be talking to Hamilton first.

From a critical perspective (such as mine), it hardly matters. The GPI is not derived from any formal framework that would identify clearly what items should be included and how they should be weighted in the index. Hence, ‘consistency with a definition’, whichever definition was chosen, would not enable detailed questions of that nature to be resolved. For this reason, Lawn’s calculation of NNP-as-*ex post*-Hicksian income in his various examples is open to the criticism that the calculations might simply turn out differently if the prices used were derived formally and made explicit.

The detour into matters of thermodynamics, which clearly concern Lawn a lot, does not seem compelling to me. The thermodynamic laws clearly *matter*, because they govern the universe, but it is not at all clear what their precise implications are for our future prospects. As long as Lawn and others associate every dollar of rising GDP with a rising volume of physical output, then we will continue to have a problem of communication here, since many dollars (and a growing share) of GDP are now derived from services produced by human capital, and disembodied goods that are non-rival in consumption, weakening the constraints imposed by entropy.

In my previous piece, I described the GPI as an ‘undisciplined’ measure, something I still believe. Many approaches to sustainability measurement, the Genuine Savings index being possibly the most well known, are derived from economic theory, which imposes its own discipline. It is true that measures of this kind have a definite Hicksian flavour, but the elements of the construction have been guided by the usage of explicit models where all the elements and assumptions are laid out for all to see. In turn, GDP, which is not a measure of economic theory but of national accounting, is guided by strict accounting identities which impose their own discipline. The creators of the GPI are not guided by any particular discipline in its construction and calculation.

And for all the seemingly abstract talk of theory, method and methodology here, these things *matter*. The GPI, as measured, shows in almost all cases a downward trend in recent decades. If the GPI is telling us something meaningful then we’re in serious trouble. The alternative position need not be that we’re *not* in trouble; it could be that the trouble we may be in is more complex than would be revealed by a detailed analysis of the GPI time series, in which case, staring at various GPI series with a magnifying glass would not be an efficient use of time. As things stand, to many critics including myself, the inverted U-shape time series of many national GPI series does not look like the breaching of a threshold, but instead as an artefact of their construction.

A detailed literature on the GPI that doesn’t eschew formal model building and the discipline that imposes? Now *that* would be genuine progress.

References

- Coyle, D. (1997), ‘The weightless economy’, *Critical Quarterly*, Vol. 39, No. 4, pp.92–98.
- Hamilton, C. (2007) ‘Measuring sustainable economic welfare’, in Atkinson, G., Dietz, S. and Neumayer, E. (Eds): *The Handbook of Sustainable Development*, Edward Elgar.
- Hamilton, K. (1994) ‘Green adjustments to GDP’, *Resources Policy*, Vol. 20, No. 3, pp.155–168.
- Harris, M. (2007) ‘On income, sustainability and the ‘microfoundations’ of the Genuine Progress Indicator’, *International Journal of Environment, Workplace and Employment*, Vol. 3, No. 2, pp.119–131.
- Hausman, D.M. (1989) ‘Economic methodology in a nutshell’, *Journal of Economic Perspectives*, Vol. 3, No. 2, Spring, 115–127.
- Lawn, P.A. (2000) *Toward Sustainable Development: An Ecological Economics Approach*, CRC Press.
- Lawn, P.A. (2003) ‘A theoretical foundation to support the Index of Sustainable Economic Welfare (ISEW), Genuine Progress Indicator (GPI), and other related indexes’, *Ecological Economics*, Vol. 44, pp.105–118.
- Leijonhufvud, A. (1973) ‘Life among the Econ’, *Western Economic Journal*, Vol. 11, No. 3, pp.327–337.

- Mäler, K-G. (2008) 'Sustainable development and resilience in ecosystems', *Environmental and Resource Economics*, Vol. 39; No. 1 pp.17–24.
- Mäler, K-G., Xepapadeas, A. and de Zeeuw, A. (2003) 'The economics of shallow lakes', *Environmental and Resource Economics*, Vol. 26, No. 4; pp.603–624.
- Nævdal, E. (2006), 'Dynamic optimisation in the presence of threshold effects when the location of the threshold is uncertain – With an application to a possible disintegration of the Western Antarctic ice sheet', *Journal of Economic Dynamics and Control*, Vol. 30, No. 7, pp.1131–1158.
- Neumayer, E. (1999) 'The ISEW: not an index of sustainable economic welfare', *Social Indicators Research*, Vol. 48, pp.77–101.
- Neumayer, E. (2000) 'On the methodology of the ISEW, GPI, and related measures: some constructive suggestions and some doubt on the threshold hypothesis', *Ecological Economics*, Vol. 34, pp.347–361.
- Randall, A. (1993) 'What practising agricultural economists really need to know about methodology', *American Journal of Agricultural Economics*, Vol. 75 (75th Anniversary Issue), pp.49–59.
- Walker, B., Holling, C.S., Carpenter, S.R. and Kinzig A. (2004) Resilience, adaptability and transformability in social–ecological systems. *Ecology and Society*, Vol. 9, No. 2, p.5. Available online at: <http://www.ecologyandsociety.org/vol9/iss2/art5>

Notes

- 1 See Leijonhufvud's entertaining 1973 piece on life among the Econ.
- 2 A long and separate discussion could be held about various ethical frameworks that could be brought to bear on discussions about sustainability. We lack space to pursue that here, and Lawn and I are both interested in the construction and interpretation of a fundamentally economic measure, so we shall stick to that.
- 3 Some identities are identities by definition, where the things on the right-hand side of the identity sign are *defined* to be the thing on the left-hand side. Accounting identities are identities by construction.
- 4 Debates about the elements included in the GPI, and the way they are weighted and aggregated, have been features of the GPI/ISEW literature (as highlighted in Harris, 2007, and see also the articles by Neumayer).
- 5 Lawn is of course aware of this, as the discussion in Lawn (2000) makes plain.