

Fairness and Tax Policy – a response to Mankiw’s proposed “Just Deserts”

Jonathan Weinstein

Kellogg School of Management, Northwestern University

December 10, 2010

I recently came across an address by Greg Mankiw, [“Spreading the Wealth Around: Reflections Inspired by Joe the Plumber.”](#) With the famous exchange between Samuel “Joe” Wurzelbacher and Barack Obama as his launching point, Mankiw suggests that the utilitarian framework¹ commonly used by economists to analyze optimal taxation conflicts with moral intuition. Mankiw argues, and I agree, that most people believe that taxation should be decided by principles of fairness: the pay people keep should be proportional to their contribution to society. He calls this “Just Deserts² Theory.” He makes a very good point this far, but I must debate some implications he draws. In fact, I believe Mankiw has identified the strength of conservative rhetoric in the public debate on tax policy, and given one answer to [“What’s the matter with Kansas?”](#), i.e. why the working class vote against their own interests. Conservatives have successfully caricatured the progressive position as placing value on equality of outcomes for its own sake, and this is a value which most Americans find distasteful. In fact, I think life in a world with equality of outcomes is fundamentally unappealing; the struggle to do better, and to be recognized fairly for achievement and productivity, is a basic human drive we would not want to lose. This is why, when conservatives convince the public of Mankiw’s basic position that the free market is the ideal, perfectly fair arbiter of Just Deserts, they win broad support for their policies, even when these policies favor the few over the many. Just as serfs once accepted that their position was allotted to them by a divine order, today’s growing inequality in wealth is considered acceptable if it is the outcome decreed by the ideal, uncorrupted

free market. Progressives must make it clear that they support the premise of fair compensation for the contributions of each individual, but dispute the notion that fairness is best achieved by an extreme *laissez-faire* version of capitalism. I'll start with some verbal arguments, then, on a slightly more technical level, will point out the flaws in Mankiw's applications of classic economic theorems before suggesting a different formal approach to fairness.

Mankiw's concluding sentence (which he is careful to qualify with "it is very possible") implies that Obama comes to his tax policy by utilitarian considerations while Joe's are based on fairness. While Mankiw is fair enough to briefly acknowledge that one could make a Just Deserts case for progressive taxation, his main argument is that fairness would favor Joe the Plumber's preferred (less progressive) tax plan over Obama's. Let's look at these claims. First, let's examine a little context surrounding the conservative movement's favorite Obama quote, "spread the wealth around." Here is the full paragraph from Obama's remarks to Joe:

"It's not that I want to punish your success. I just want to make sure that everybody who is behind you, that they've got a chance at success, too... My attitude is that if the economy's good for folks from the bottom up, it's gonna be good for everybody. If you've got a plumbing business, you're gonna be better off [...] if you've got a whole bunch of customers who can afford to hire you, and right now everybody's so pinched that business is bad for everybody and I think when you spread the wealth around, it's good for everybody."

Giving everyone a "chance at success" is a notion of fairness. In fact, it is the classic conservative notion of fairness: equality of opportunity, not equality of outcome. The subsequent sentences are not a statement of values at all, but are a claim in the realm of positive economics: that a tax code that favors the less wealthy will stimulate the economy and be good for everyone. Obama would have a lot of support in this claim, but true or false, it is not a utilitarian justification. A utilitarian would say something like "If you make a lot of money, the last \$10,000 isn't doing you much good; it would make a much bigger difference to poor people." "It is very possible" that Obama thinks this way, but he certainly didn't say so to Joe³.

In fact, I think Mankiw is quite right that most people on both ends of the spectrum form their beliefs about taxes based on notions of fairness. But what exactly does fairness demand? Here's the view of Warren Buffett, quoted in Obama's book *The Audacity of Hope*:

“I happen to have a talent for allocating capital. But my ability to use that talent is completely dependent on the society I was born into. If I’d been born into a tribe of hunters, this talent of mine would be pretty worthless...but I was lucky enough to be born in a time and place where society values my talent, and gave me a good education to develop that talent, and set up the laws and the financial system to let me do what I love doing - and make a lot of money doing it. The least I can do is help pay for all that.”

Now, everyone is dependent on a civil society for the dollars they earn, and even for the existence of dollars. Buffett seems to be making the further case that financiers like him are more dependent, as a percentage of income. Buffett does a nice job of explaining this case intuitively, and I will later elaborate on this and suggest some ideas of how we might formalize the additional dependence. Mankiw gives a bit of a nod to the notion that some ultra-rich don’t deserve their full income, but only cites the most obvious cases, “the CEO who pads the board with his cronies and the banker whose firm survives only by virtue of a government bailout.” What about legitimately successful traders like Buffett? He would include himself as deserving higher taxes, but Mankiw seems not to. Mankiw exempts the non-corrupt ultra-rich, citing Spielberg, Jobs, Letterman and Rowling as among those who produce a large value and deserve to keep it. Now, it is true that large-scale entertainers, inventors, and financiers can create a huge value. It may sound fair for them to keep their marginal value, or perhaps have it taxed at the same rate as working-class salaries - to Mankiw, and to conservatives in general, but, as Buffett would recommend, we should question this premise. Fairness, of course, ultimately depends on value judgments, but nonetheless we can elucidate these judgments through analysis. Mankiw outlines such an analysis, arguing for the conservative position that everyone deserves their free-market outcome, which will equal their marginal product, but I noticed that he seriously misuses the classic economic theorems he cites. This is a key passage (p. 295):

“From the perspective of classic liberalism, it is natural to presume that any individual, or group of individuals, should be allowed to leave the large society to live on their own and form smaller communities. They exercise this right if they feel their contributions are insufficiently rewarded - that is, if they can do better on their own. This freedom ensures that the resulting allocation of resources will be in what game theorists call the core...Debreu and Scarf proved that, as the number of players gets

large, the core of such games converge to the competitive equilibria...For any other allocation, some group will exercise their right to leave because they are not getting their just deserts.”

I have always thought that the conclusion of *Atlas Shrugged*, in which the noble captains of industry go to their own secluded valley to be free of the oppressive liberal state, is a hilarious, unintentional, *reductio ad absurdum* against the entire position of the book. I mean, really, how exactly are financiers, railroad managers, entertainers or authors, who profit immensely from living in a large industrial society, going to achieve anywhere close to those profits by moving to a tiny community of elites? Now Mankiw’s argument seems to say that it is an actual mathematical theorem that Atlas should shrug, and will profit by doing so, if he receives anything less than his full value under *laissez-faire*. This is a bit like a [proof that pi is four](#); it must be wrong, and we just have to figure out how. What essential aspect of the Core Equivalence Theorem has been ignored?

In an exchange economy, a competitive (or Walrasian) equilibrium is an allocation that can occur when each player optimizes his or her behavior, taking some vector of prices as given. The Core Equivalence Theorem says that, in the limit where we make many copies of the original economy, any non-Walrasian allocation will be outside the core⁴. The key word here is “copies.” The quantity that must become large for the theorem to apply is not the “number of players,” but the number of copies of each kind of player. Intuitively, this is essential in order to force each player to act as a price-taker who cannot use monopoly power to extract rents. The proof depends crucially on the fact that if someone has too large a share, his or her clones will defect along with the rest of the population. Mankiw has used the theorem in precisely the cases where it fails: when high earners extract rents thanks to their uniqueness. Besides which, anyone whose output is easily scalable, including inventors, entertainers, and financiers, is able to create value in proportion to the size of the economy. Fairness would seem to dictate that they owe a greater proportion of their income to the upkeep of society than those in non-scalable professions such as hairdresser, teacher, etc.

There is a similar problem with a result Mankiw cites in the previous paragraph, where he says “...it is also a standard result that in a competitive equilibrium, the factors of production are paid the value of their marginal product.” This result also requires that there be many duplicates of every type of person - in Mas-Colell et. al.

the result is proved for economies with an infinite multiplicity of each type, and it is noted that the result typically fails when this is not the case⁵. In practice, think of infinite as meaning large enough that no one person can affect the price of any good - this is clearly not the case for the iPad! Without even thinking of the technical results, we should question whether it is always possible for everyone to earn their marginal product. Don't we live in a world with huge complementarities, since at least the days of Adam Smith and the pin factory? Doesn't that mean that the sum of our marginal products may be far larger than our total product?

Furthermore, even when marginal-value allocation is possible, it may not match a reasonable notion of fairness. Consider the stylized example of a single owner of capital and a large set of interchangeable workers with decreasing marginal value. The workers are hired until their marginal value is zero. The marginal value of the owner, who is indispensable, is the full product of the industry. Is this fair? Reflect that if the workers can unionize and be treated as a single player, now the marginal product of the union is equal to that of the owner, and they will negotiate for the surplus on roughly equal terms. (They can no longer both make their marginal products, which will add to double the actual surplus.) An advocate of the notion that market freedom leads to Just Deserts must answer the question "Whose freedom⁶?" That is, is it a more fundamental freedom for workers to be free to organize and bargain collectively, or for capitalists to be free to negotiate with individual workers as they see fit? Market outcomes depend not just on productivity, but also on bargaining power. A worker's Just Deserts presumably do not depend on whether he lives in a union-friendly regime, but his outcome depends enormously. Whichever outcome seems fair to you, the other cannot be, so a blanket assertion that "Free market outcomes are fair" falls apart; it is either incomplete or self-contradictory⁷.

If marginal value is an unworkable notion of fairness, what might be a better one? A classic idea in cooperative game theory which it would be fascinating to apply here is [Shapley value](#). This is related to marginal value. It is computed by the thought experiment of letting individuals enter the economy in a random order. Your marginal value will then depend on who is there when you enter. Your Shapley value is simply your average marginal value across all permutations. Shapley famously showed that this is the unique method of allocating surplus which satisfies some very basic principles of fairness⁸. Note also that it always gives a unique outcome, while the core frequently either fails to exist or has a huge range of possible outcomes⁹.

Shapley value is commonly used in arbitration, but I have never seen it discussed in the context of tax policy. What sort of results would it give? I find this an enormously interesting question, to which I can only offer some preliminary thoughts here. As a first approximation, someone in a highly scalable profession would keep roughly half their income, since they enter the game with, on average, half the population present. (See a more carefully worked out example in the appendix.) There are many possible adjustments to this estimate; for one, if the inventor or entertainer is extracting rents from network effects and they are not actually much better than a replacement, their Shapley value might be much less than half their income. On the other hand, someone in a non-scalable profession creates roughly the same value regardless of the size of society, so they would keep more of their income. Whether these considerations reflect fairness is, of course, ultimately a value judgment, but a 50% top marginal tax rate is well within the historical range, so such an outcome would not be radical.

The great intellectual advances that illuminated the enormous benefits of the free market, starting with Adam Smith and continuing into the 20th century, have long since been assimilated into our political discourse. The danger is that in some circles the lessons have been learned just a bit too well. The free market then becomes a 21st-century deity whose dictates are perfectly fair and should not be questioned, lest its manna of prosperity cease to rain down upon us. Warning about this is, of course, unnecessary for economists, who, whatever their political stripe, understand perfectly the limits of core equivalence and welfare theorems. Keeping any nuance is very difficult when intellectual advances are distilled for a larger population, so responsible academics always have to be very careful in how they discuss the practical impact of abstract results.

1 Appendix

I'll give here the Shapley values and core for a simple production economy, with some commentary. This is basically Exercise 259.3 and its follow-ups in Osborne and Rubinstein's *A Course in Game Theory*; I felt like working it out and thinking through the economic consequences for my own benefit, and I hope readers' benefit as well.

Let there be one owner of capital and a large number of workers who we can

approximate as a continuum of unit mass. The value of a coalition is 0 if it does not have the capitalist, otherwise if there is a mass x of workers present, the value is $f(x)$ for a concave function f , normalized so that $f(0) = 0, f(1) = 1$. Since, in a random permutation, the mass of labor present when the capitalist enters has uniform distribution, the Shapley value for the owner is the average of f across this uniform distribution, that is, $c = \int_{x=0}^1 f(x)dx$. By concavity of f , this satisfies $.5 \leq c \leq f(.5)$. That is, the owner gets more than half the surplus, but less than the product of half the workers. Labor divides the remaining surplus $l = 1 - c$ equally, so that $1 - f(.5) \leq l \leq .5$. If f happens to be linear, all of these bounds hold with equality, so surplus is equally divided between capital and labor.

The core can be shown to consist of all allocations where each worker makes at most the marginal product of the last worker, $f'(1)$. The key coalitions you have to worry about consist of the owner and all but one worker; it turns out that concavity implies if you satisfy these coalitions, you are in the core. Therefore, for core allocations the owner gets at least $1 - f'(1)$. (Note that it is always in the core for the owner to get everything; here is a simple example where core equivalence is violated, despite a large number of players, if one player is unique.) The core is very harsh towards replaceable workers; if $f'(1) = 0$, the unique core outcome is that the workers get nothing, although they are necessary to production. This sort of example illustrates why traditionally the core is seen as a descriptive solution concept - an outcome driven by bargaining power - while the Shapley value is more of a normative concept, driven by fairness. If someone engages in a productive activity but is replaceable, it does seem natural that this would diminish the value he keeps - but by how much? Zero seems unfair; even if the very last worker adds almost no value, I don't think many would agree that we should then pay them all in peanuts. Like the core, Shapley value changes if labor is treated as a single player - but not as drastically. It turns out to be sort of a compromise between the fully unionized outcome, and the harsh, non-unionized, core outcome.

If Shapley value is more fair than outcomes in an unfettered free market, how might we move closer to it, using a system which will of course be founded on the free market, but mediated by tax and labor policies? There is a literature on "non-cooperative foundations" which deals with questions such as this, but apparently there is no good known non-cooperative path to Shapley value. Even if no mechanism leads precisely to Shapley value, though, it could still provide guideposts as to what general

direction policy might move.

Notes

¹The version of utilitarianism he discusses says that we decide on our policy by maximizing the sum of individual utilities, where utility is concave, meaning that an extra dollar held by a rich person adds less to social welfare than one held by a poor person. If total wealth were constant, social welfare would then be maximized by a completely equal distribution. Utilitarianism does not actually advocate that all wealth be shared equally, however, because it acknowledges the “leaky bucket” principle which says that transfers create inefficiencies by diminishing incentives. However, the leaky bucket is the only rationale that utilitarianism allows for unequal outcomes; there is no notion that the highly productive may deserve their wealth. More general versions of utilitarianism combine individual utilities via a non-linear function, or an arbitrary weighted sum. The best justification for maximizing the simple sum (or average) of utilities is Rawls’ famous veil of ignorance; that is, if you were about to be randomly assigned to any position in society, you would want society to maximize average utility.

²In a previous version, I erroneously stated that “deserts” was misspelled. It is an uncommon meaning of the word, “that which is deserved,” and while pronounced like an after-dinner sweet is spelled like the Sahara. Many thanks to a commenter for pointing this out.

³I would be unjust not to include some additional comments I found from Obama, that do refer to decreasing marginal utility of wealth: “...you can derive as much pleasure from a Picasso hanging from a museum as from one that’s hanging in your den.” (*The Audacity of Hope*, p. 193.) But this wasn’t the argument he made to Joe. My best read of Obama is that his views are primarily based on fairness, with some additional utilitarian motivation. My larger point is independent of how Obama himself may think: it hardly follows from advocacy of progressive taxation that one’s motivation is utilitarian rather than fairness-based.

⁴As Mankiw described, the core consists of precisely those allocations for which no subpopulation can leave the group and do better on their own. For details on the core equivalence theorem, see for instance the PhD textbook by Mas-Colell, Whinston and Green (MWG), pp. 652-659.

⁵MWG, 670-672

⁶This is the title of a wonderful book by linguist George Lakoff, who is famous for describing how our view of the world is shaped by our use of language. In “Whose Freedom?” he describes how the battle between the conservative and progressive meanings of “freedom” is central to national politics.

⁷Depending on whether the intended meaning is some particular outcomes, which have been underspecified, or all free-market outcomes.

⁸See the Wikipedia link, or, for instance, Osborne and Rubinstein’s *A Course in Game Theory*.

⁹Shapley value and core are generally unrelated; however, if the production function is convex, meaning that anyone’s marginal product is always higher when more people, of any type, are added to the economy, then the Shapley value is in the core (MWG 683). Convexity tends not to hold

everywhere, though; when the people being added are of the same type as you, in many cases your marginal value decreases. When convexity fails, the core may be empty.