

Post-print (accepted manuscript).

Acerbi, A., & Sacco, P. L. (2023). Self-Interest, prosociality, and the moral cognition of markets: A comparative analysis of the Theory of Moral Sentiments and the Wealth of Nations. *Rationality and Society*, 0(0). <https://doi.org/10.1177/10434631231170460>.

First published online April 18, 2023.

**Self-interest, prosociality, and the moral cognition of markets:
A comparative analysis of *The Theory of Moral Sentiments*
and *The Wealth of Nations***

Running head: Adam Smith and moral cognition

Abstract

In this paper, we perform a text analysis of Adam Smith's two books, the Theory of Moral Sentiments and the Wealth of Nations, to better characterize their highly disputed differences in terms of moral cognition. In particular, given that Smith's ideas are still very cited and influential in the current scholarly debate on moral cognition, we are interested in understanding whether a text analysis would unveil a semantic structure that is in line with a dual process theory interpretation or, alternatively, with a neuro-emergent cognition one. We find that, despite that the intellectual koine in which Smith's thought was originally embedded would be more in line with a dual process theory approach, the analysis reveals a better consonance with the neuro-emergent cognition approach. This opens new and interesting perspectives in future research on the moral cognition of market interactions in a Smithian tradition of thought.

Keywords: markets, moral cognition, self-interest, sympathy, dual process theory, neuro-emergent cognition approach.

JEL Codes: B12, D91, Z13

Our elaborate market exchange system owes its existence not to our calculating brain or insatiable self-centeredness, but rather to our sophisticated and nuanced human sociality and to the inherent rationality built into our emotions.

Aysar (2020), p. 89.

1. Introduction

The moral cognition of market exchange is far from being a problem of exclusive interest to cognitive scientists, social psychologists and moral philosophers. Our assumptions and expectations about the intentions and motives of others when entering a transaction that involves strangers have a clear impact on how we behave in such interaction, and in turn the interaction's outcome will feed back on our beliefs about what others think and feel and how they will act consequently. This complex process builds on the interplay between cognitive and emotional aspects and may even imply some stipulation about constituent features of human nature.

A clear illustration in this regard is the normative value that is often assigned to the assumption that people tend to be essentially self-interested in market interactions. This assumption derives from the schematic characterization of human nature that has come to be universally known as *homo economicus* (Persky, 1995). Once widely adopted and legitimized, the norm of self-interest becomes self-fulfilling and the adoption of alternative, other-regarding rules of conduct can only emerge as the result of intentional, highly coordinated agency, generally in the pursuit of prosocial goals (Grant and Patil, 2012). The normative force of the self-interest assumption can be so strong that acting prosocially to support causes in which one does not have a personal stake may elicit moral outrage and reproach by third party observers, and people tend to internalize such attitude, expecting to be judged negatively should they engage in such behavior without a valid personal motive (Ratner and Miller, 2001), because (only) “interest will not lie” (Long, 1991). Consequently, according to the self-interest norm perspective, even when people are willing to engage in prosocial behaviors such as donating to a charity, they need to *rationalize* this as a form of self-interest (for instance, by receiving a gift in exchange for their donation) to avoid cognitive dissonance (Simpson et al, 2006). It follows that, as a rule, people tend to selectively reconstruct the hidden motives for selfless actions in self-interested terms, whereas they do not question whether behind apparently self-interested actions there

could be selfless motives, thereby ‘naturalizing’ self-interest as the default moral cognition orientation (Critcher and Dunning, 2011). Not only cynical beliefs about the inner motives of other tend to be customarily held and persist, but they also actively resist disconfirmation. Such normative judgment is further socially corroborated by the widespread belief that self-interested individuals are cognitively superior in that self-interested conduct is more adaptive than an other-regarding one. However, large sample, inter-cultural evidence shows that the opposite is true: self-interested individuals tend to underperform in cognitive ability and competency tasks (Stavrova and Ehlebracht, 2019).

Despite the widespread fascination with the idea that self-interest need be more adaptive than prosociality, it must be noted that a normative characterization of self-interest is far from universal and typical of Western culture (Miller, 2001). There are for instance human cultures where socially favorable side effects of action are more likely to be judged as intentional, and thus motivationally salient, than socially unfavorable ones (Robbins et al, 2017). Moreover, there is an increasingly solid and consistent evidence that humans have deep-seeded prosocial instincts such as those for fairness and justice, which are already well developed in childhood, although in culturally nuanced forms (Eisenberg and Mussen, 1989), and are shared also by other nonhuman primates (Brosnan, 2013). It is therefore not surprising that also for those interactions in which the self-interest norm is generally presumed to be particularly compelling, such as those between CEOs and their corporate boards, the mediating role of norms of reciprocity and fairness as bounds to self-interest has to be advocated to make space for actual welfare improving corporate practices (Bosse and Phillips, 2016). Even when the issue of the moral cognition of markets is addressed from an extreme position such as that implied by a self-interest norm, there is no way to escape the necessity to account for richer characterizations of human nature where prosocial motivations and dispositions play a relevant role.

The moral cognition of markets therefore calls for an understanding of the relative prevalence of self- vs. other-regarding orientations in different social situations, as modulated by their respective cognitive and emotional elements (Moll et al, 2005). The effect of market interaction on prosocial orientations is however ambiguous. If on the one side markets promote reliability and other-regarding behaviors as desirable characteristics to be sought in potential market interaction partners, on the other hand they emphasize the self-centered motives that may lead people to extract as much value as possible from any interaction, at the expense of the partner (Zaki et al, 2021). In the history of economic thought, no figure epitomizes this tension better than Adam Smith. Smith’s two major works, *The Theory of Moral Sentiments* (1759, 1990) (TMS) and the *Wealth of Nations* (1776) (WoN) have long puzzled historians of economic thought for their apparently divergent focus: towards prosocial orientations in TMS, as encapsulated in the Smithian notion of sympathy, and towards self-

interest in WoN, as crystallized in the so-called Invisible Hand Theorem. Such ambiguity is at the root of the so-called Adam Smith problem, whose current relevance goes beyond the appraisal of the internal consistency of Smith's oeuvre, and extends to contemporary moral cognition debates (Montes, 2003, 2019).

The reason behind the surprising modernity of Adam Smith in the moral cognition debate is that he does not maintain a fixed view of human nature, and develops a nuanced distinction between natural sentiments, that are shaped by cultural selection processes, and moral sentiments, that are developed as the result of a complex acculturation process (Schliesser, 2011; Pack and Schliesser, 2018). Actually, Smith goes as far as to anticipate the idea that moral behaviors might be shared by nonhuman animals as well (Brosnan, 2011), further emphasizing the open ended, non-essentialist character of his idea of morality. This conceptual flexibility, and the intrinsic ambiguity of Smith's oeuvre when considered as a whole, possibly incomplete philosophical system, has paved the way to many, even mutually contradictory appropriations of his thought. One key building block of Smith's thought is the idea of the impartial spectator, which is congenial to his notion of sympathy and can be seen as instrumental to the definition of two different instances: second-person versus third-person sympathy, which can be taken in turn as a possible conceptual foundation of direct vs. indirect reciprocity (Haig, 2011). Third-person effects, in particular, are an often overlooked but very important facet of prosocial behavior (van der Iest et al, 2011).

From the perspective of a dual process theory of moral judgment (Greene et al, 2001, 2004), for instance, the impartial spectator can be 'naturally' regarded as an internal regulator that overrules the impulsive, emotionally driven judgments dictated by human passions in favor of cognitively controlled, carefully pondered ones (Meardon and Ortmann, 1996), and this makes of Smith a direct precursor of behavioral economics in its overarching normative focus on impulsive and emotionally charged behavior as a source of bias and bounded rationality that needs to be purposefully neutralized (Ashraf et al, 2005) – an interpretation that provides however a poor match to Smith's notion of rationality (Dixon and Wilson, 2014), and that need not be invoked to support a price-theoretic rather than decision-theoretic interpretation of rational choice (Hudik, 2019).

In this paper, we develop a relatively uncommon approach to the implications of Adam Smith's thought for the contemporary moral cognition debate, by undertaking a direct textual analysis of Smith's two major works. Our approach provides a precise characterization of the structural differences between TMS and WoN in the coverage of a number of semantic categories of major importance for moral cognition. Our analysis is centered on the computation of the relative frequency of topic-relevant lexica across the texts of the two works for a select number of topics ranging from family to work, from religion to money. In this way, we are able to show that, indeed, TMS and WoN

not only have, as already remarked in the ‘Adam Smith problem’ debate, a different thematic focus, but they are, literally, covering very different topical spectra in different ways. As pointed out by Vernon Smith (2010), one can characterize TMS as a theory of personal social exchange, and WoM as a theory of impersonal market exchange so that, in fact, such disjunction makes good conceptual sense. However, a consequence of the disjunction is that the lack of extensive conceptual overlaps between the two works makes it possible to reinterpret Smith’s system in radically different ways, by selectively emphasizing certain elements over others, or by forcing certain interpretations of part of the system to give meaning to others, with very different implications in terms of moral cognition. In particular, it is interesting to assess how such ambiguity plays out in the current debate between the already mentioned dual process-based approach to moral cognition, which has been largely endorsed by the currently mainstream behavioral economics research, and the alternative paradigm that conceptualizes choice and its related moral concerns in terms of neuro-emergent cognition (Moll and Schulkin, 2009; Molnar-Szakacs, 2011; Zak, 2011b; Fogassi, 2011), according to which the sympathetic perspective of the impartial spectator, which can be seen as a personification of moral cognition itself (Bréban and Gilardone, 2020), and indirectly as the dispositional foundation of social order (Hardin, 2013), can be operationalized by the mirror neuron system (Khalil, 2011). We will show how, despite that Smith’s intellectual roots would in principle resonate more with a dual process approach to moral cognition in which rational decision making calls for close cognitive control and even inhibition of automatic emotional response (e.g. Buon et al, 2016), his system of thought, once analyzed in terms of the semantic spectra of his two major works, seems to be more in line with the neuro-emergent cognition view, and thus with an idea of emotion as a fundamental *enabler* of moral cognition (e.g. Moll et al, 2007). This is a further, surprising confirmation that Smith’s system of thought still offers us an unvaluable, open-ended conceptual scaffolding to frame current debates and to formulate new discriminating hypotheses and experiments to better understand human moral cognition.

The remainder of the paper is organized as follows. Section 2 briefly reviews some aspects of the ‘Adam Smith problem’ debate that provide a convenient way to characterize the fruitful ambiguity (from the viewpoint of the current debate) of Smith’s conception of moral cognition. Section 3 introduces our textual analysis and presents its methodology and our main research hypotheses. Section 4 presents the results of the textual analysis. A discussion of the results and their implications for the broader debate on moral cognition is provided in Section 5. Section 6 concludes.

2. *The fruitful ambiguity of the Adam Smith problem*

The most straightforward way to interpret the difference in focus between TMS and WoN is, quite simply, that Smith changed his mind along the way, initially postulating that prosociality played a key role in human matters, and at some point coming to the conclusion that self-interest was a more parsimonious explanation that was not incompatible with the possibility of socially beneficial forms of coordination and cooperation by means of market exchange, and this was indeed the prevailing thesis in nineteenth century scholarship. However, the idea that there could be ‘two Adam Smiths’ was made increasingly puzzling by the fact that Smith kept on revising the TMS until his death, which seems incompatible with the conviction that he had discarded that leg of his system of thought (Otteson, 2000). On the other hand, the thesis that, as suggested by Vernon Smith (2010), the two works refer to different spheres of human affairs, namely personal social exchange and impersonal market exchange, respectively, still leaves open the problem that there is no reason why sympathy as administered by the impartial spectator should not apply to market transactions as well, unless one wants to consider the possibility that sympathy and more generally prosocial orientations are only sociobiologically supported for in-group relations whereas self-interest (or even negative reciprocity) generally applies to out-group relations (Greene, 2013). Vernon Smith’s idea resonates with Brown’s (2009) discussion of two distinctive forms of agency that are characteristic of each of the two works, namely moral agency in TMS, and economic agency in WoN. Paganelli (2008) argues instead that Smith’s treatment of self-interest in TMS is in fact more favorable, and has more solid social foundations, than in WoN, so that the apparent contradiction is basically due to an incomplete understanding and a biased interpretation of TMS, and WoN should rather be read as a problematization of the role of self-interest in human affairs, rather than as a statement of universal relevance. This view resonates with Emma Rothschild’s (1994) discussion of the notion of the Invisible Hand as an essentially un-Smithian idea, which Smith mainly used as an ironic, paradoxical example. However, other interpretations regard the Invisible Hand as a central tenet of Smith’s thought, in strict continuity with the line of argument developed in TMS. In this perspective, the Invisible Hand is the metaphor of a superior principle of social organization that is conceptually derived from the Enlightenment Deist idea of the watchmaker supervising the perfect mechanical functioning of the universe (Evensky, 1987), or that even lends itself to Providentialist readings in a theological (Hill, 2001; Oslington, 2012) or stoic (Jones, 2010) tradition.

The surprising centrality that Smith’s thought still has in contemporary debates is a direct product of this ambiguity, and in particular of the fact that the ‘Adam Smith problem’, which was a debate that sprung among the German readers of Smith in the second half on the nineteenth century while at the time in England Smith’s thought was largely taken for granted and uncontroversial, had the effect

of focusing scholarly attention on the many still under-researched layers of his thought – a process that since then has kept on engaging one generation of scholars after another (Tribe, 2008). In other words, the relevance of Smith’s thought seems to stem from the fact that his system can be closed in many possible ways, by projecting upon it different interpretive frameworks with markedly different implications in terms of moral cognition. However, reading Smith a-historically as a mere canvas for contemporary debates inevitably leads to flawed and partial interpretations of his thought (Tribe, 1999).

Even this very brief and partial synthesis of the debate shows how Smith’s thought allows the dialectical coexistence of almost opposite interpretations, each of which can be legitimately rooted in a purposeful selection and interpretation of some parts of his argument. A relatively crude but thorough way to overcome this difficulty is to avoid selective readings and to analyze each of the two books as a text that covers a certain semantic spectrum as characterized by its vocabulary and by the relative frequency of use of the terms that it contains. An analysis that proceeds from similar premises to ours, and that offers complementary insights, is that of Graafland and Wells (2021). Moreover, Martin (2021) carries out another textual analysis of Smith’s work to address a different question, namely Smith’s attitude toward the workers and the poor.

Therefore, the best way of exploiting the fruitful ambiguity of Smith’s system to reflect on moral cognition from a contemporary perspective while at the same time not losing sight of the original context of his reflection is to analyze the two books by means of a comparative characterization of their semantic spectra. How are TMS and WoN different in terms of their relative coverage of topics relevant for moral cognition in the respective bodies of text? How important are the emotional vs. cognitive spheres in the semantic universe of each of the books, and what is the relative importance of positive vs. negative emotions? These are the questions to which we turn our attention now.

3. Textual analysis: methodology and research hypotheses

We analyzed the complete texts of TMS and WoN by means of the LIWC tool (<http://liwc.wpengine.com>). The digital texts were retrieved from https://en.wikisource.org/wiki/The_Theory_of_Moral_Sentiments and <http://www.gutenberg.org/cache/epub/3300/pg3300.txt> respectively. LIWC carries out a computerized text analysis by means of categories (for a detailed presentation of the conceptual development of the instrument and of its psychometric properties see Pennebaker et al, 2007). A LIWC category is a list of words, or words stems, that are associated semantically with a

psychologically relevant dimension. One such category is, for instance, sadness (**sad**), whose associated entries include *abandon** (the stem for words such as *abandoned*, *abandonment*, etcetera), *cry*, *suffer*, and others, for a total of around 100 entries. We can thus identify within a text the frequency of entries associated to a particular category, giving us an indicator of the salience of such category within the semantic space of the text. These categories span several areas of interest from the moral cognition viewpoint, including the emotional sphere, the cognitive sphere, the social sphere, and a number of areas of human affairs of special relevance.

The total list of the LIWC categories that we considered for the analysis of the two texts, as well as examples of words for each category are presented in Table 1.

Sphere	LIWC category	Full name	Example words
Emotional	affect	affect (general emotional tone)	happy, cried, abandon
	negemo	negative emotion	hurt, ugly, nasty
	posemo	positive emotion	love, nice, sweet
	anx	anxiety	worried, fearful, nervous
	anger	anger	hate, kill, annoyed
	sad	sadness	crying, grief, sad
Cognitive	cogmech	cognitive mechanisms	cause, know, ought
	insight	insight	think, know, consider
	cause	causation	because, effect, hence
	discrep	discrepancies	should, would, could
	tent	tentativeness	maybe, perhaps, guess
	inhib	inhibition	block, constrain, stop
	incl	inclusive	and, with, include
	excl	exclusive	but, without, exclude
	certain	certainty	always, never
Social	social	social	mate, talk, they, child
	family	family	daughter, husband, aunt
	friends	friends	buddy, friend, neighbor
	humans	humans	adult, baby, boy
Human affairs	work	work	job, major, resource

	leisure	leisure	cook, chat, movie
	achiev	achievement	earn, hero, win
	home	home	apartment, kitchen, family
	money	money	audit, cash, owe
	relig	religion	altar, church, mosque
	death	death	bury, coffin, kill

Table 1. List of LIWC categories used for the analysis, with example of words for each category.

Frequencies of words associated to each category are compared in TMS and WoN. We also use binary logistic regression models, where the binary dependent variable is the book (TMS vs. WoN) and the independent variables are the frequencies of words in each category. All p values of the regression models, for each comparison, are reported in Appendix 1. We run regression models with a binomial distribution and with a Poisson distribution, to take into account the possibility that our independent variable may be zero-inflated, i.e., many words present in the LIWC categories may not be found in the books. In the text, we report the (more conservative) p value of the Poisson regression.

It is worth remarking that, as in our analysis we are not working with samples from a stochastic data generating process but with texts that are the result of an intentional process of literary composition by a human writer, the differences in the relative frequencies of terms for a given semantic category cannot strictly speaking be interpreted as significant or not in a rigorous statistical sense. The intentional choice by Smith of a certain word in a certain passage in the text could make a big difference in terms of its overall interpretation, even if occurring only once. However, such occurrence would likely not be ‘significant’ in statistical terms. Therefore, even when differences in relative frequencies of certain semantic categories are not ‘significant’, the fact that some difference actually exists between the two texts remains of interest and deserves to be considered and discussed. With these caveats, the use of p values nonetheless gives us important insights as to the relative frequencies of semantic categories in the two texts and will be reported accordingly.

What could we expect from this analysis in terms of insight into moral cognition issues? As discussed in the previous section, a key difference between TMS and WoN is that the former focuses upon the social sphere and on moral agency, whereas the latter upon the economic sphere and agency. If one assumes that, however incomplete, Smith’s system in its most definitive form builds on both books and that therefore sympathy can be considered as a foundational category of Smith’s approach to moral cognition, we need to consider how the emotional vs. cognitive spheres are related according

to alternative theoretical perspectives. If we want to interpret Smith's thinking in a dual process theory perspective, we therefore expect that in the context in which rational behavior is more salient and normative, namely in the economic sphere analyzed in the WoN, there should be a strong emphasis on the capacity of the cognitive sphere to override the emotional one, as rational behavior calls for an effective control of impulsive, emotionally driven behavior. On the other hand, as TMS is essentially focusing upon the moral and emotional dimension, here the tension between the cognitive and the emotional should be less salient, and we should observe a clear prevalence of the emotional dimension over the cognitive one. Consequently, in WoN the semantic sphere of cognition should be more represented than the emotional one, and vice versa for TMS.

On the other hand, taking the alternative perspective of neuro-emergent cognition, we rather expect that, if Smith's thinking is more consistent with this interpretation, the emotional and cognitive spheres would be deeply intertwined in terms of an emergent rather than purposive notion of 'rationality' (Macy, 1997), with no obvious semantic prevalence of cognition neither in TMS nor in WoN. Since in the perspective of neuro-emergent cognition the key focus is not upon the domain of self-interested action but rather upon that of the deep motives of social interaction, we would expect that both the cognitive and emotional semantic spheres are comparatively well represented in TMS with respect to WoN, in the context of a general prevalence of the representation of emotion in TMS with respect to WoN. From this alternative perspective, WoN has a subordinate role with respect to TMS as an exploration of the complexity of human motives, and the emphasis on self-interest might also be read as a sort of counterfactual exercise, a sort of extreme stress test of the resilience of a system of social exchange that is generally based upon prosocial orientation (Hausken, 1996). We then expect that, in this perspective, the semantic space of WoN would be characterized by a focus on specific contexts of human affairs such as work, money, and achievement, as domains of practical application of Smith's system, rather than on basic moral cognition topics.

Of course, proving that Smith's system can be more naturally read in terms of a specific moral cognition perspective rather than another does not imply that any of these are correct in themselves or not, but simply that they make more or less sense in a Smithian perspective. However, given the continuing relevance of Smith's thought in the contemporary scholarly debate, and the wide recognition that Smith's views on moral cognition still remain a source of interesting and generative ideas, checking to what extent a given approach is consistent with Smithian thinking also bears some interesting implications as to which research agenda would be worthier to pursue as a further step in the Smithian tradition (Paganelli, 2011).

4. Results

We begin the presentation of our results with the analysis of the comparative frequency of emotional categories in TMS vs. WoN. In this regard, TMS has twice as many terms related to the emotional semantic sphere than WoN (i.e., 8% of terms in the **affect** category for TMS as compared to 4% for WoN, $p < 0.001$). In Figure 1, as well as in all subsequent figures, we report the overall relative frequency of that category in TMS vs. WoN, and the list of the 20 terms that contribute the most to that category in each of the two texts.

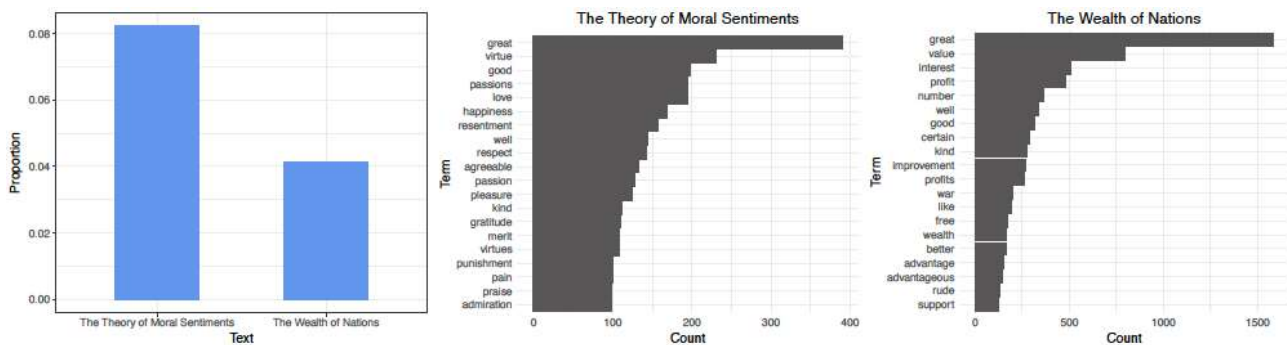


Figure 1. Relative frequency of affect-related terms in TMS vs. WoN.

We can now focus on the two sub-categories of **posemo** and **negemo**. First, it is interesting to stress that, as far as general emotional tone is concerned, WoN is emotionally more positive than TMS, as 74% of all the emotional terms are positive (calculated as the proportion between **affect** and **posemo**) in the former whereas only 63% are in the latter. Notice that the fact that in both texts there are more terms related to positive emotions than to negative ones is not meaningful as positive terms are in general more abundant than negative ones. What matters is the relative proportion between the two texts.

As shown in Figure 2, the frequency of both positive and negative emotions is larger in TMS as compared to WoN, but, consistently with the previous finding, the gap is much narrower for positive emotion ($p = 0.06$) than it is for negative ones ($p < 0.001$), since the frequency of negative emotions in TMS is almost three times as much as in WoN. Notice how the most frequent **negemo** term in WoN is ‘number’, an artefact of the presence of the numb* word stem in the respective lexicon. This unwanted feature is however not impinging upon our results, as excluding the term ‘number’ would make the incidence of **negemo** terms even lower in WoN, further strengthening our result.

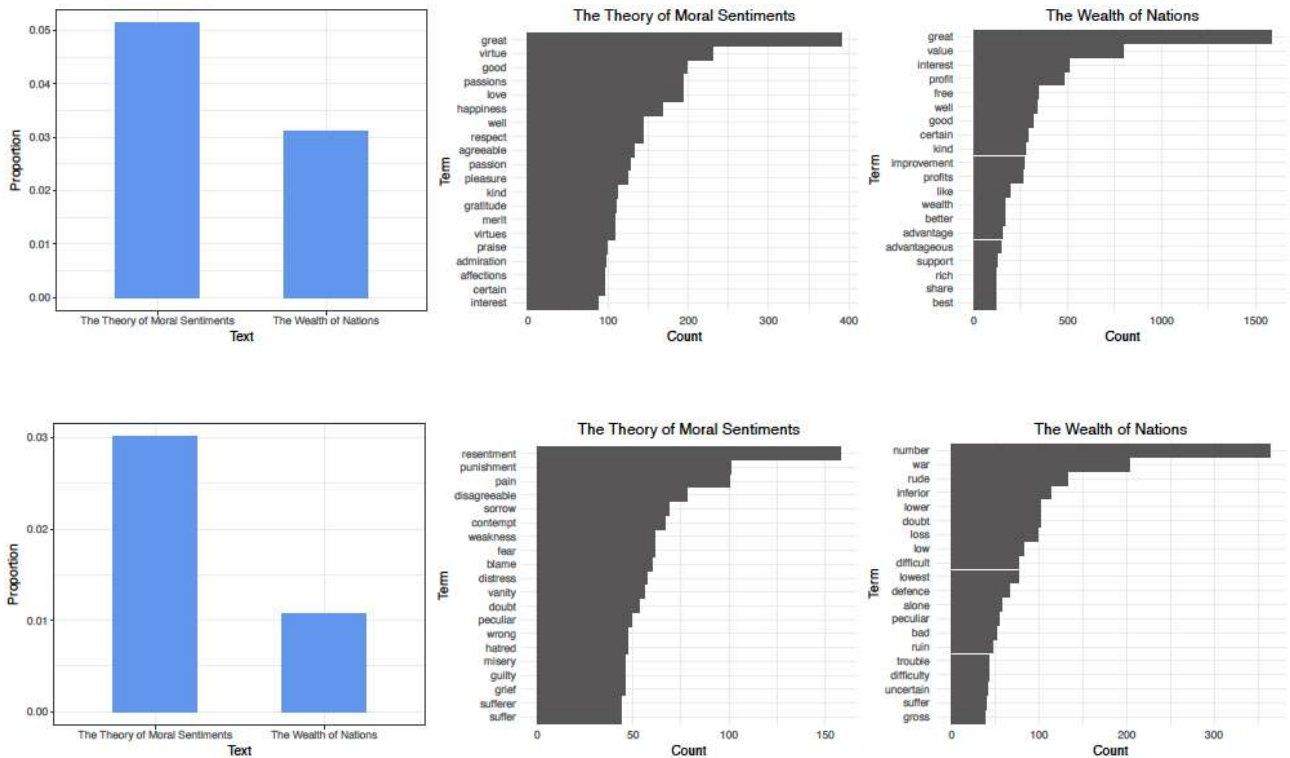


Figure 2. Positive (top) and negative (bottom) emotion related terms in TMS vs. WoN.

The relative frequencies for the emotion sub-categories **anx**, **anger** and **sad** are not shown for brevity. For anxiety, the gap between TMS and WoN is even more considerable, with a fivefold frequency in the former with respect to the latter ($p < 0.01$); for anger, the gap is still very large but about fourfold ($p < 0.01$); and for sadness is slightly more than double ($p < 0.05$). Therefore, also in each sub-dimension of the emotional sphere, negative emotions are much more prevalent in TMS than in WoN.

Coming now to the cognitive sphere, the overall **cogmech** score presents, quite interestingly, a slight prevalence in frequency in TMS over WoN, even if not ‘significant’ ($p = 0.57$), as shown in Figure 3.

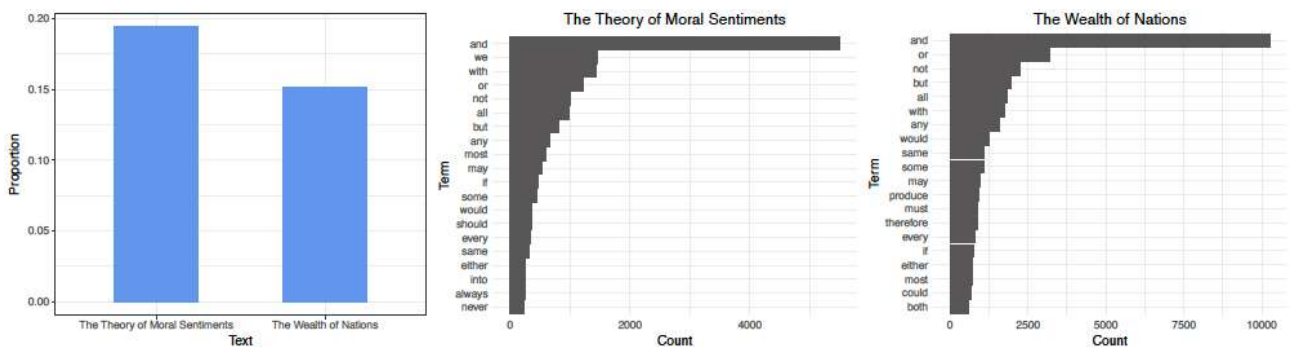


Figure 3. Relative frequency of cognition-related terms in TMS vs. WoN.

Coming to the specific cognitive sub-categories, there are two where WoN marginally prevails over TMS: causation and inhibition, although in the first case the frequencies are almost equal ($p = 0.95$), and in the second the difference is still very slight ($p = 0.69$), as shown in Figure 4.

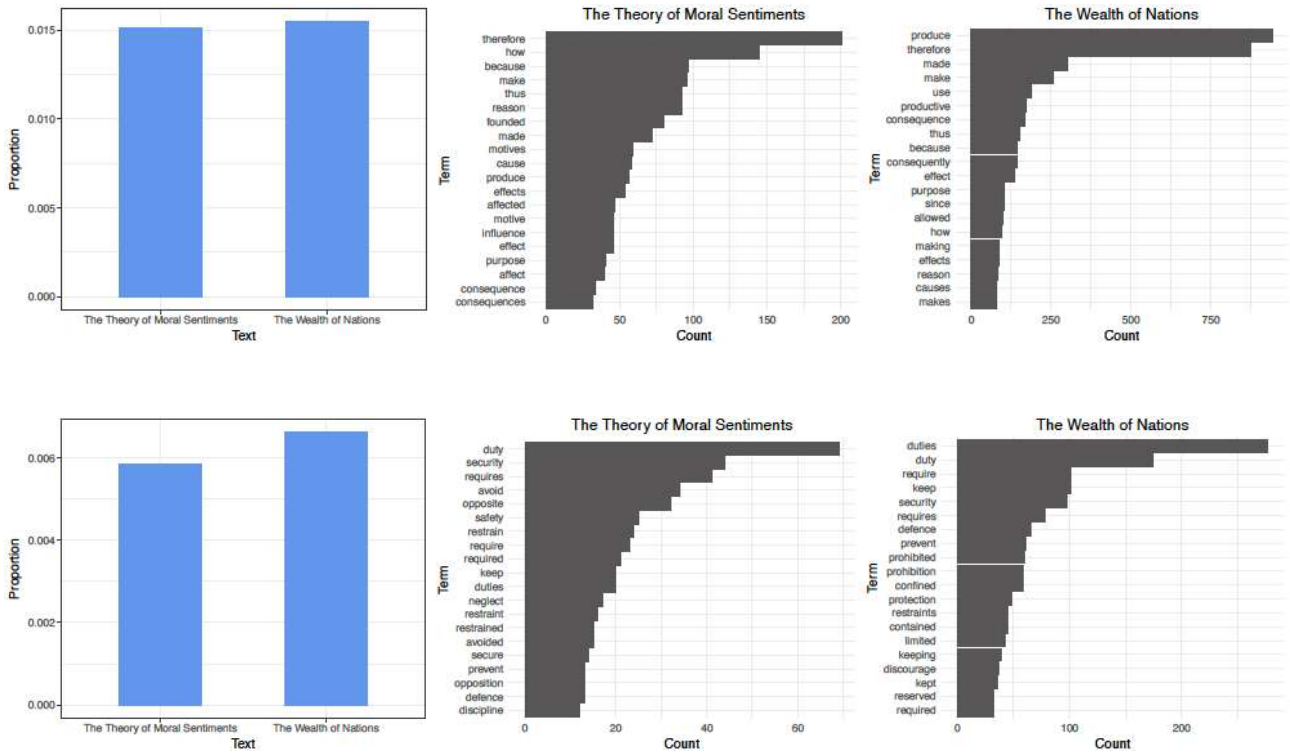
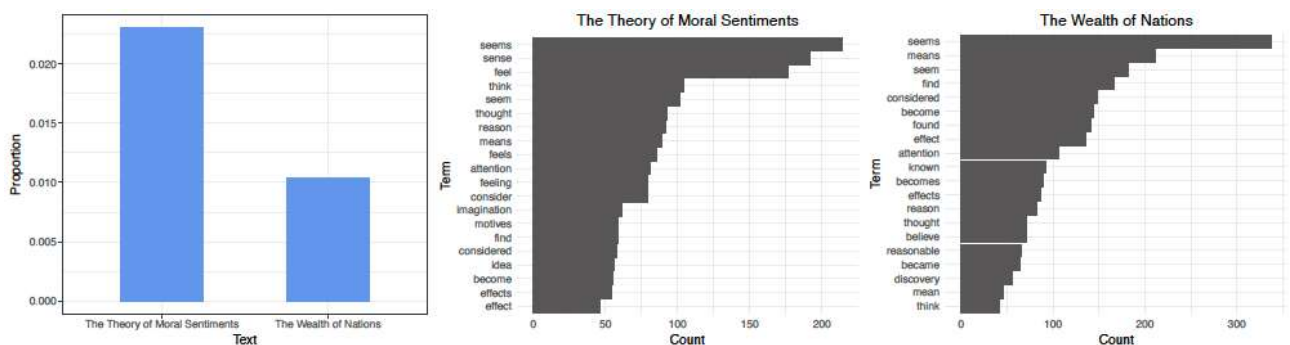


Figure 4. Relative frequency of *cause* (top) and *inhib* (bottom) in TMS vs. WoN.

On the other hand, for all other cognitive sub-categories, TMS prevails over WoN. In particular, there is a slight prevalence for exclusion (**excl**, $p = 0.93$), some prevalence for tentativeness (**tentat**, $p = 0.86$), discrepancies (**discrep**, $p = 0.80$) and certainty (**certain**, $p = 0.78$), and a relatively large difference for inclusion (**incl**, almost double, $p = 0.73$) and especially **insight** (more than double, $p < 0.05$); the latter two are shown in Figure 5.



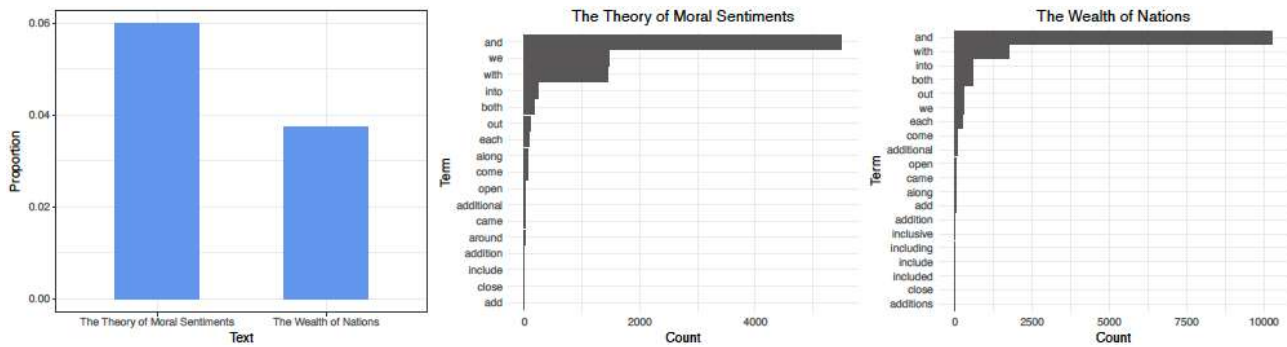
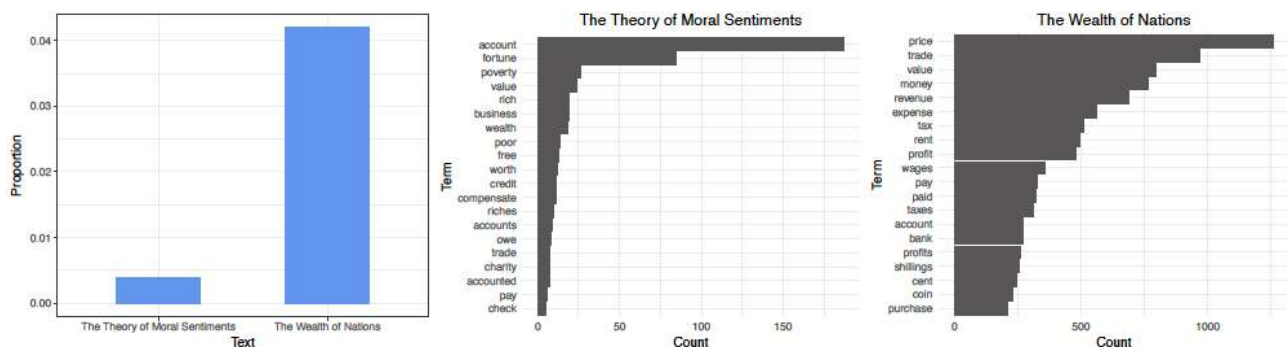


Figure 5. Relative frequency of *insight* (top) and *incl* (bottom) in TMS vs. WoN.

Coming to the social dimension, the general sociality category **social**, as it could be expected, is twice more represented in terms of frequency in TMS with respect to WoN ($p = 0.14$). And the gap remains wide for the specific social sub-categories: it is biggest for **friends** (more than threefold, $p = 0.14$), but also large for **family** ($p = 0.41$) and **humans** (roughly twice, $p = 0.38$).

Finally, considering the spheres of human affairs, we find a clear separation between the two books. We have three spheres of affairs in which WoN strongly prevails: **money** (tenfold, $p < 0.001$), **work** (fourfold, $p < 0.001$) and, maybe more surprisingly, **home** (sixfold, $p = 0.11$). WoN also slightly prevails upon TMS in **leisure** ($p = 0.55$), whereas, possibly again surprisingly, they are practically equivalent for achievement (**achiev**), with a marginal prevalence of TMS ($p = 0.94$). The latter is instead clearly prevalent for religion (**relig**, $p = 0.29$) and especially **death** (more than twofold, $p = 0.33$). In Figure 6, we show the results for money, home, achievement and death.



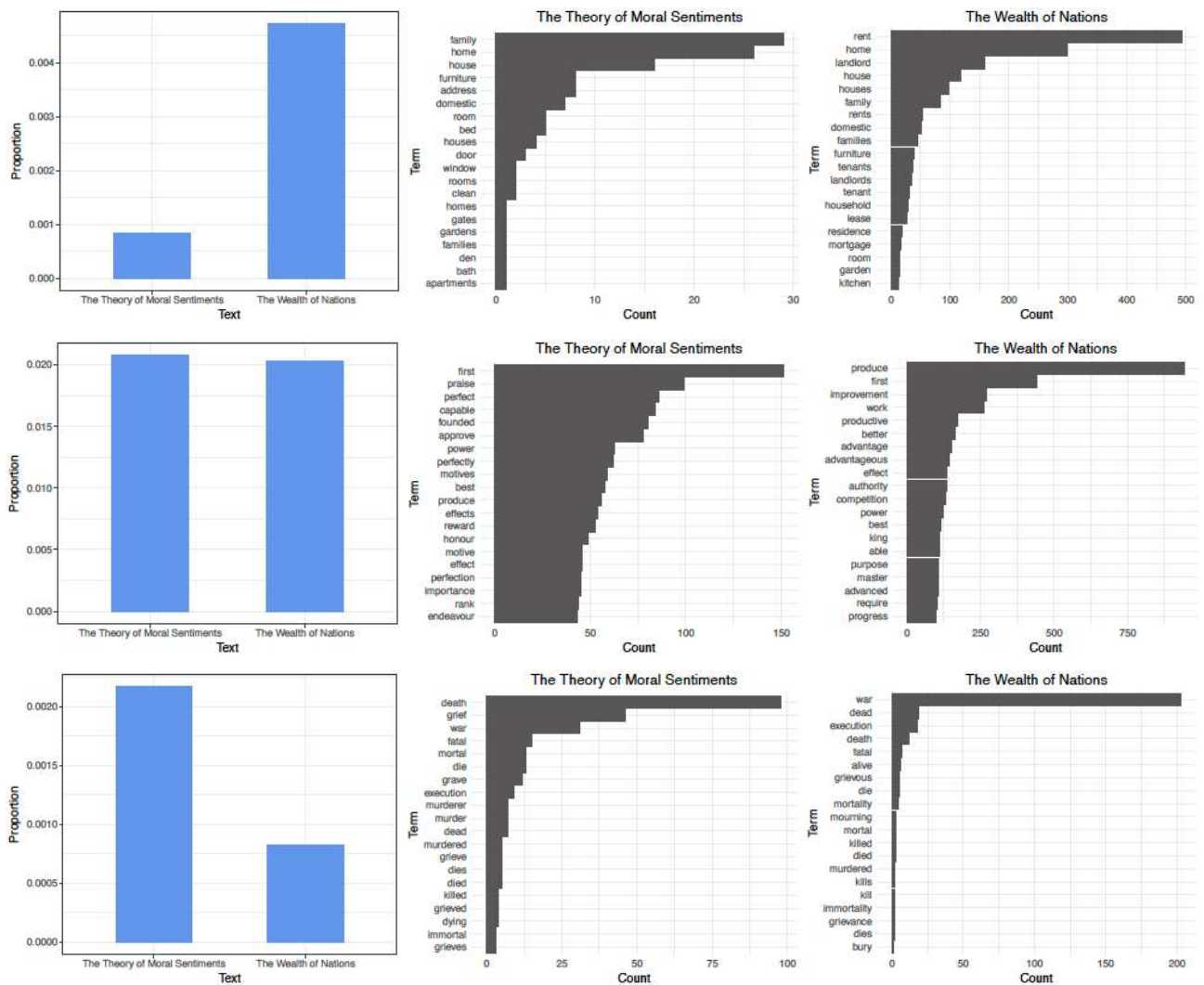


Figure 6. Relative frequency of (from top to bottom) *money*, *home*, *achiev** and *death* in TMS vs. WoN.

Overall, we can then conclude that TMS clearly prevails over WoN in terms of semantic space devoted to emotion, both in terms of general affect and of positive and negative emotion, but particularly so on the negative emotional side. At the same time, TMS also slightly prevails in the cognitive sphere, with the marginal exception of inhibition and causation, whereas all the other cognitive categories show a varying level of prevalence of TMS. TMS is also clearly prevailing in the social sphere, both in general terms and with respect to specific sub-categories. The only domain where we find a strong prevalence of WoN is in certain human affairs categories, namely money, work and home, and less strongly leisure, whereas the two books are practically equivalent concerning achievement, and TMS prevails in religion and death.

5. Discussion

In her recent paper on the relevance of Adam Smith's thinking for contemporary moral cognition debates, Songhorian (2021) argues that Smith provides a convenient conceptual framework to address some of the currently most pressing issues, and in particular the integration of the emotional and cognitive dimensions. However, she also notices that Smith was well-aware of the intrinsic limitations of both sympathy and the impartial spectator as socio-behavioral regulatory mechanisms for the development of fair judgment and of a mature moral conscience. There is therefore the necessity of a reflective moral agency conditional on social and cultural circumstances (Jackson, 2021), informed by solid contextual knowledge (Mueller, 2021), and possibly robust with respect to hedonic moral hypocrisy (Lindenberg et al, 2018). This implies that any extreme, universalist interpretation of Smith's thinking, both in terms of sympathy as a self-sufficient foundation for prosocial behavior, and in terms of the role of the impartial spectator as the cognitive locus of control of emotional impulses and irrational exuberance that is conducive to rational choice, are likely to be far-fetched. To what extent our results align with these conclusions?

First, our results show that there is no clear prevalence of the cognitive semantic sphere in the WoN with respect to the TMS and rather, and possibly surprisingly for someone, references to cognition are slightly prevalent in the latter. The only categories where cognition is preferentially represented in the WoN are causation and inhibition, and while this is interesting as it is somewhat resonating with a dual process theory interpretation, so that in the domain of the economic, causal reasoning and inhibition of impulse clearly hint at cognitive planning and control functions, it must also be acknowledged that the differences between the two texts in this regard are too marginal, and below *ex ante* expectations, to make this truly meaningful. TMS is richer semantically than WoN in the emotional sphere *and* basically equivalent in the cognitive sphere, thus bringing some support to the alternative interpretation rooted in the neuro-emergent cognition literature that the cognitive and emotional dimensions are equally functional and jointly essential in determining an adaptive response to the social environment, and in enabling moral cognition. It is especially interesting how TMS is considerably semantically richer than WoN in the treatment of negative emotions, which are typically associated to social situations of competition and conflict, and that both books cover equivalently the category of achievement, that would intuitively be much more strongly associated to the economic agency domain of the WoN than to the moral agency domain of the TMS. What these results seem to imply is that Smith essentially addresses and systematizes the issues related to competition and conflict in the emerging market society of his time within the conceptual framework of sympathy and

the impartial spectator, with all their limitations, rather than in the (possibly counterfactual) framework of self-interest – a result that largely agrees with the meta-analysis of secondary literature on Smith's thought carried out by Hühn and Dierksmeier (2016). At the same time, our results also clearly show that, in line with expectations, Smith's treatment of human sociality is essentially contained in TMS and only marginally in WoN. The latter is, ultimately, mostly related to the thorough analysis of specific domains of human affairs, and especially of those related to economic activity, such as the sphere of personal and family security that comes with economic success and affluence (Rasmussen, 2006).

Our analysis seems therefore to suggest that, in a sense, the legacy of Smithian thought for the moral cognition debate seems to mainly reside in the TMS, and that consequently it is especially the dimension of moral agency (and moral imagination) that needs to be considered in this regard, also in the light of the social constructivist approach that informs Smith's philosophy of science (Hühn, 2019). Interestingly, the results obtained by Graafland and Wells (2021) through a different text-mining analysis and a methodology that is substantially complementary to ours yield very similar implications. Yet another concurrent interpretation emerges from the exegetical analysis of Smith's works by Bevan and Werhane (2015). As far as moral cognition is concerned (but clearly not regarding many other key issues and topics of great relevance for economics), WoN may be considered as an interesting repository of case studies, possibly analyzed by means of counterfactual arguments such as testing whether the market system would still work when – as a paradoxical, extreme instance – people are basically self-interested, but always within the broader context of sympathy and the impartial spectator as defining features of the human. For instance, Smith never contemplates the possibility that the pursuit of self-interest would justify callousness or psychopathy, for example by causing intentional harm to other humans to increase one's profits (Kurz, 2016; Pack and Schliesser, 2018). Rather, the key of the argument is that in the context of market exchange the structure of incentives is such that even a self-interested individual would find it convenient, in their own interest, to act as if they were prosocial (Caldas et al, 2007), that is, by providing their customers with the best possible products to beat the competition – although the complexity of human nature and the strategic incentives from wealth accumulation call for prudence in taking human prosociality too much for granted (Paganelli and Schumacher, 2019). It is therefore rather typical to observe 'moderate' prosociality orientations such that 'generous' acts imply the transfer of relatively lower-value resources (Harrell, 2021).

In this regard, the moral agency focus of the TMS suggests that, contrary to the common interpretation, the construction of a market society is not, and especially should not, be regulated by a self-interested norm, but rather by an other-regarding norm that developmentally orientates people

in their social interaction attitudes throughout the life course (Zahn-Waxler et al, 1995), and is mediated by psychological satisfaction (Glaze, 2017), an implication that conforms to the most recent research on the role of norm psychology in the emergence of prosociality (House et al, 2020). This raises further concern about the potential damage of uncritical, self-fulfilling adoption of a norm of self-interest in human moral cognition (Kish-Gephart et al, 2014).

Our results also resonate with recent criticism of dual process theories in terms of their failure to properly account for the role of emotion regulation in moral cognition (Helion and Pizarro, 2015), and in particular of acknowledging the increasingly documented complementarity of emotion and cognition in co-determining moral judgments (Helion and Ochsner, 2018) and providing a source of internal moral constraints (Stringham, 2011) – a perspective that presents many points of contact with the TMS notion of moral agency (Zak, 2011a).

6. Conclusions

Recent advances in neuroscientific research show how moral decisions are the result of a complex functional integration of distributed, heterogeneous networks that simultaneously engage moral, social, empathic, and unconstrained cognition. Moral cognition moreover presents intriguing parallels with the mind wandering activity that is typical of the Default Mode Network, that is, the brain's default pattern of activity (Bzdok et al, 2015). The idea that the impulsive, automatic neurobiological processes associated to emotion are a potential threat to rational decision-making that needs constant control and that has to be overtaken by slower, more reflective, cognitively-driven mental processes to guarantee an adaptive response to environmental conditions is increasingly at odds with our understanding of how human moral cognition is rooted in deep cortical mechanisms that provide a very complex and adaptively sophisticated regulation of our choice in social situations (Moll and de Oliveira-Souza, 2007), and with the increasingly recognized, key role of emotions in the reciprocal attunement of subjects engaging in prosocial interaction (Capraro and Halpern, 2019). It is probably in this tradition that Smith's legacy to the current moral cognition debate finds its most appropriate context. As already remarked, Smith was himself convinced, very much in line with the neuro-emergent cognition approach, that moral judgments were shared also by nonhuman animals, even if the basic difference between humans and animals seems to be the reflective nature of such judgments. Whereas nonhuman animals are only sensitive to moral dilemmas that concern themselves, humans are also responsive to dilemmas that concern third parties – once again reaffirming the fruitfulness of the impartial spectator in characterizing some key aspects of human moral cognition (Brosnan, 2011).

In principle, an intellectual figure so deeply rooted in Enlightenment thinking such as Adam Smith would be expected to be much more in line with a tradition of thought that essentializes rational deliberation as an exquisitely cognitive, emotionally un-charged activity such as that of dual process theory, but in fact what our analysis shows, and what further emerges from the wider discussion of the literature, is that Smith's thought presents some surprising features of an ante litteram evolutionary thinking (Boulding, 1991; Hodgson, 1998; Klein, 2003), that ultimately fits rather naturally within the current consilience of socio-anthropological, psychological and neurobiological research on human moral cognition (Schliesser, 2011). But until very recently, many economists were still reading Smith in outdated evolutionary terms that were still regarding narrow self-interest as a necessary or at least likely outcome of Darwinian selection and invoked Smithian sympathy and more generally moral emotions as a possible escape route from socially inferior solutions of moral dilemmas (Frank, 2011). In fact, Smith's thought is much better served by a dialogue with more recent evolutionary biology perspectives which assign to moral emotions a much more profound and crucial role. And this is probably the reason why we keep on engaging with Smith's two masterworks, whereas we have not been paying attention anymore to so many other books, often by key figures of the history of economic thought, which are much closer to us in historical time.

References

Ashraf N, Camerer CF, Loewenstein G (2005) Adam Smith, behavioral economist. *Journal of Economic Perspectives* 19, 131-145.

Aysar R (2020) *The evolutionary origin of markets. How evolution, psychology, and biology have shaped the economy*. Routledge, London.

Bevan D, Werhane P (2015) The inexorable sociality of commerce: The individual and others in Adam Smith. *Journal of Business Ethics* 127, 327-335.

Bosse DA, Phillips RA (2016) Agency theory and bounded self-interest. *Academy of Management Review* 41, 276-297.

Boulding KE (1991) What is evolutionary economics? *Journal of Evolutionary Economics* 1, 9-17.

Bréban L, Gilardone M (2020) A missing touch of Adam Smith in Amartya Sen's account of public reasoning: The man within for the man without. *Cambridge Journal of Economics* 44, 257-283.

Brosnan SF (2011) An evolutionary perspective on morality. *Journal of Economic Behavior and Organization* 77, 23-30.

Brosnan SF (2013) Justice- and fairness-related behaviors in nonhuman primates. *Proceedings of the National Academy of Sciences* 110, 10416-10423.

Brown V (2009) Agency and discourse: Revisiting the Adam Smith problem. In Young JT (ed), *The Elgar Companion to Adam Smith*, Elgar, Cheltenham.

Buon M, Seara-Cardoso A, Viding E (2016) Why (and how) should we study the interplay between emotional arousal, Theory of Mind, and inhibitory control to understand moral cognition? *Psychonomic Bulletin & Review* 23, 1660-1680.

Bzdok D, Gross D, Eickhoff SB (2015) The neurobiology of moral cognition: Relation to theory of mind, empathy, and mind-wandering. In Clausen J, Levy N (eds), *Handbook of neuroethics*, Springer, Dordrecht, 127-148.

Caldas JC, Narciso Costa A, Burns TS (2007) Rethinking economics: The potential contribution of the classics. *Cambridge Journal of Economics* 31, 25-40.

Capraro V, Halpern JY (2019) Translucent players: Explaining cooperative behavior in social dilemmas. *Rationality and Society* 31, 371-408.

Critcher CR, Dunning D (2011) No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology* 47, 1207-1213.

Dixon W, Wilson D (2014) Political economy and the social disciplines: The modern life of *Das Adam Smith Problem*. *Cambridge Journal of Economics* 38, 623-641.

Eisenberg N, Mussen PH (1989) *The roots of prosocial behavior in children*. Cambridge University Press, Cambridge.

Evensky J (1987) The two voices of Adam Smith: Moral philosopher and social critic. *History of Political Economy* 19, 447-468.

Fogassi L (2011) The mirror neuron system: How cognitive functions emerge from motor organization. *Journal of Economic Behavior and Organization* 77, 66-75.

Frank RH (2011) The strategic role of the emotions. *Emotion Review* 3, 252-254.

Glaze S (2017) Adam Smith and William James on the psychological basis of progress. *Cambridge Journal of Economics* 41, 349-365.

Graafland J, Wells TR (2021) In Adam Smith's own words: The role of virtues in the relationship between free market economies and societal flourishing. A semantic network data-mining approach. *Journal of Business Ethics* 172, 31-42.

Grant A, Patil SV (2012) Challenging the norm of self-interest: Minority influence and transitions to helping norms in work units. *Academy of Management Review* 37, 547-568.

Greene J (2013) *Moral tribes. Emotion, reason and the gap between us and them*. Penguin, London.

Greene J, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105-2108.

Greene J, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389-400.

Haig D (2011) Sympathy with Adam Smith and reflexions on self. *Journal of Economic Behavior and Organization* 77, 4-13.

Hardin R (2013) The priority of social order. *Rationality and Society* 25, 407-421.

Harrell A (2021) How can I help you? Multiple resource availability promotes generosity with low-value (but not high-value) resources. *Rationality and Society* 33, 341-362.

Hausken K (1996) Self-interest and sympathy in economic behaviour. *International Journal of Social Economics* 23, 4-24.

Helion C, Ochsner KN (2018) The role of emotion regulation in moral judgment. *Neuroethics* 11, 297-308.

Helion C, Pizarro DA (2015) Beyond dual-processes: The interplay of reason and emotion in moral judgment. In Clausen J, Levy N (eds), *Handbook of neuroethics*, Springer, Dordrecht, 109-125.

Hill L (2001) The hidden theology of Adam Smith. *European Journal of the History of Economic Thought* 8, 1-29.

Hodgson GM (1998) Evolutionary economics. In *Handbook of economic methodology*, Edward Elgar, Cheltenham, 160-164.

House BR, Kanngiesser P, Barrett HC, Broesch T, Cebiglu S, Crittenden AN, Erut A, Lew-Levy S, Sebastian-Enesco C, Smith AM, Yilmaz S, Silk JB (2020) Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behavior* 4, 36-44.

Hudik M (2019) Two interpretations of the rational choice theory and the relevance of behavioral critique. *Rationality and Society* 31, 464-489.

Hühn MP (2019) Adam Smith's philosophy of science: Economics as moral imagination. *Journal of Business Ethics* 155, 1-15.

Hühn MP, Dierksmeier C (2016) Will the real A. Smith please stand up! *Journal of Business Ethics* 136, 119-132.

- Jackson BR (2021) Custom and moral variability: An interpretation of part V of Smith's *Theory of Moral Sentiments*. *Journal of Economic Behavior and Organization* 184, 804-812.
- Jones HB (2010) Marcus Aurelius, the stoic ethic, and Adam Smith. *Journal of Business Ethics* 95, 89-96.
- Khalil EL (2011) The mirror neuron paradox: How far is *understanding* from *mimicking*? *Journal of Economic Behavior and Organization* 77, 86-96.
- Kish-Gephart J, Detert J, Klebe Treviño L, Baker V, Martin S (2014) Situational moral disengagement: Can the effects of self-interest be mitigated? *Journal of Business Ethics* 125, 267-285.
- Klein S (2003) The natural roots of capitalism and its virtues and values. *Journal of Business Ethics* 45, 387-401.
- Kurz HD (2016) Adam Smith on markets, competition and violations of natural liberty. *Cambridge Journal of Economics* 40, 615-638.
- Lindenberg S, Steg L, Milovanovic M, Schipper A (2018) Moral hypocrisy and the hedonic shift: A goal-framing approach. *Rationality and Society* 30, 393-419.
- Long D (1991) Taking interest seriously. *Rationality and Society* 3, 343-364.
- Macy MW (1997) Identity, interest and emerging rationality. An evolutionary synthesis. *Rationality and Society* 9, 427-448.
- Martin C (2021) Adam Smith and the poor: A textual analysis. *Journal of Economic Behavior and Organization* 184, 837-849.
- Meardon SJ, Ortmann A (1996) Self-command in Adam Smith's *Theory of Moral Sentiments*. A game-theoretic interpretation. *Rationality and Society* 8, 57-80.
- Miller DT (1999) The norm of self-interest. *American Psychologist* 54, 1053-1060.
- Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J (2005) The neural basis of human social cognition. *Nature Reviews Neuroscience* 6, 799-809.
- Moll J, de Oliveira-Souza R (2007) Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive Sciences* 11, 319-321.
- Moll J, de Oliveira-Souza R, Garrido GJ, Bramati IE, Caparelli-Daquer EMA, Paiva MLMF, Zahn R, Grafman J (2007) The self as a moral agent: Linking the neural bases of social agency and moral sensitivity. *Social Neuroscience* 2, 336-352.
- Moll J, Schulkin J (2009) Social attachment and aversion in human moral cognition. *Neuroscience and Biobehavioral Reviews* 33, 456-465.
- Molnar-Szakacs I (2011) From actions to empathy and morality – A neural perspective. *Journal of Economic Behavior and Organization* 77, 76-85.

Montes L (2003) *Das Adam Smith Problem*: Its origins, the stages of the current debate, and one implication for our understanding of sympathy. *Journal of the History of Economic Thought* 25, 63-90.

Montes L (2019) Adam Smith's foundational idea of sympathetic persuasion. *Cambridge Journal of Economics* 43, 1-15.

Mueller PD (2021) Adam Smith on moral judgment: Why people tend to make better judgments within liberal institutions. *Journal of Economic Behavior and Organization* 184, 813-825.

Oslington P (2012) God and the market: Adam Smith's Invisible Hand. *Journal of Business Ethics* 108, 429-438.

Otteson JR (2000) The recurring 'Adam Smith problem'. *History of Philosophy Quarterly* 17, 51-74.

Pack SJ, Schliesser E (2018) Adam Smith, natural movement and physics. *Cambridge Journal of Economics* 42, 505-521.

Paganelli MP (2008) The Adam Smith problem in reverse: Self-interest in the Wealth of Nations and the Theory of Moral Sentiments. *History of Political Economy* 40, 365-382.

Paganelli MP (2011) The same face of the two Smiths: Adam Smith and Vernon Smith. *Journal of Economic Behavior and Organization* 78, 246-255.

Paganelli MP, Schumacher R (2019) Do not take peace for granted: Adam Smith's warning on the relation between commerce and war. *Cambridge Journal of Economics* 43, 785-797.

Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ (2007) The development and psychometric properties of LIWC 2007. University of Texas at Austin and University of Auckland, <https://www.liwc.net/LIWC2007LanguageManual.pdf> (accessed December 27, 2021).

Persky J (1995) The ethology of homo economicus. *Journal of Economic Perspectives* 9, 221-231.

Rasmussen DC (2006) Does 'bettering our condition' really make us better off? Adam Smith on progress and happiness. *American Political Science Review* 100, 309-318.

Ratner RK, Miller DT (2001) The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology* 81, 5-16.

Robbins E, Shepard J, Rochat P (2017) Variations in judgments of intentional action and moral evaluation across eight cultures. *Cognition* 164, 22-30.

Rothschild E (1994) Adam Smith and the Invisible Hand. *American Economic Review Papers and Proceedings* 84, 319-322.

Schliesser E (2011) Reading Adam Smith after Darwin: On the evolution of propensities, institutions, and sentiments. *Journal of Economic Behavior and Organization* 77, 14-22.

Simpson B, Irwin K, Lawrence P (2006) Does a 'norm of self-interest' discourage prosocial behavior? Rationality and quid pro quo in charitable giving. *Social Psychology Quarterly* 69, 296-306.

Smith VL (2010) What would Adam Smith think? *Journal of Economic Behavior and Organization* 73, 83-86.

Songhorian S (2021) Adam Smith's relevance for contemporary moral cognition. *Philosophical Psychology*, <https://www.tandfonline.com/doi/full/10.1080/09515089.2021.2014442>.

Stavrova O, Ehlebracht D (2019) The cynical genius illusion: Exploring and debunking lay beliefs about cynicism and competence. *Personality and Social Psychology Bulletin* 45, 254-269.

Stringham EP (2011) Embracing morals in economics: The role of internal moral constraints in a market economy. *Journal of Economic Behavior and Organization* 78, 98-109.

Tribe K (1999) Adam Smith: Critical theorist? *Journal of Economic Literature* 37, 609-632.

Tribe K (2008) 'Das Adam Smith problem' and the origins of modern Smith scholarship. *History of European Ideas* 34, 514-525.

van der Iest H, Dijkstra J, Stokman FN (2011) Not 'just the two of us': Third party externalities of social dilemmas. *Rationality and Society* 23, 347-370.

Zahn-Waxler C, Cole PM, Welsh JD, Fox NA (1995) Psychophysiological correlates of empathy and prosocial behaviors in preschool children with behavior problems. *Development and Psychopathology* 7, 27-48.

Zak P (2011a) The physiology of moral sentiments. *Journal of Economic Behavior and Organization* 77, 53-65.

Zak P (2011b) Moral markets. *Journal of Economic Behavior and Organization* 77, 212-233.

Zaki J, Neumann E, Baltiansky D (2021) Market cognition: How exchange norms alter social experience. *Current Directions in Psychological Science* 30, 236-241.

Appendix 1

LIWC category	p value (binomial)	p value (Poisson)
affect	< 0.001	< 0.001
negemo	< 0.001	< 0.001
posemo	< 0.05	0.06
anx	< 0.001	< 0.01
anger	< 0.001	< 0.01
sad	< 0.01	< 0.05
cogmech	0.44	0.57
insight	< 0.001	< 0.05
cause	0.93	0.95
discrep	0.73	0.80
tent	0.81	0.86
inhib	0.58	0.69
incl	0.63	0.73
excl	0.90	0.93
certain	0.70	0.78
social	< 0.05	0.14
family	0.25	0.41
friends	0.06	0.14
humans	0.24	0.38
work	< 0.001	< 0.001
leisure	0.40	0.55
achiev	0.91	0.94
home	< 0.05	0.11
money	< 0.001	< 0.001
relig	0.17	0.29
death	0.22	0.33

Table A1: P values of the binary logistic regression models used to compare LIWC categories in TMS vs. WoN. The binomial distribution is customarily used, whereas the Poisson distribution is used to account for zero-inflated samples.