

VU Research Portal

Social Norms Offer Explanation for Inconsistent Effects of Incentives on Prosocial Behavior

Graf, Caroline; Merz, Eva-Maria; Suanet, Bianca; Wiepking, Pamala

2021

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Graf, C., Merz, E-M., Suanet, B., & Wiepking, P. (2021). *Social Norms Offer Explanation for Inconsistent Effects of Incentives on Prosocial Behavior*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Social Norms Offer Explanation for Inconsistent Effects of Incentives on Prosocial Behavior

Caroline Graf^{*,†}, Eva-Maria Merz^{*,†}, Bianca Suanet[‡]
and Pamala Wiepking^{§,*,¶}

Abstract

Incentives have surprisingly inconsistent effects when it comes to encouraging people to behave prosocially. Classical economic theory, according to which a specific behavior becomes more prevalent when it is rewarded, struggles to explain why incentives sometimes backfire. More recent theories therefore posit a reputational cost offsetting the benefits of receiving an incentive – yet unexplained effects of incentives remain, for instance across incentive types and countries. We propose that social norms can offer an explanation for these inconsistencies. Ultimately, social norms determine the reputational costs or benefits resulting from a given behavior, and thus variation in the effect of incentives may reflect variation in norms. We implemented a formal model of prosocial behavior integrating social norms, which we empirically tested on the real-world prosocial behavior of blood donation. Blood donation is essential for many life-saving medical procedures, but also presents an ideal testing ground for our theory: Various incentive policies for blood donors exist across countries, enabling a comparative approach. Our preregistered analyses reveal that social norms can indeed account for the varying effects of financial and time incentives on individual-level blood donation behavior across 28 European countries. Incentives are associated with higher levels of prosociality when norms regarding the incentive are more positive. The results indicate that social norms play an important role in explaining the relationship between incentives and prosocial behavior. More generally, our approach highlights the potential of integrating theory from across the economic and behavioral sciences to generate novel insights, with tangible consequences for policy-making.

Keywords: social norms | incentives | prosocial behavior | blood donation

*Center for Philanthropic Studies, Dept. of Sociology, Vrije Universiteit Amsterdam

†Dept. of Donor Medicine, Sanquin Research

‡Dept. of Sociology, Vrije Universiteit Amsterdam

§Lilly Family School of Philanthropy, Indiana University–Purdue University Indianapolis

¶We are grateful to Wim de Kort and all blood donation experts from across Europe who shared their knowledge about incentives for donors in their countries with us. Moreover, we thank Mark Ottoni-Wilhelm, Mario Macis, Bianca Beersma, Joris Schröder and David Schultner for thoughtful comments on earlier versions of this draft. Correspondence concerning this working paper should be addressed to Caroline Graf (c.graf@vu.nl) or Eva-Maria Merz (e.m.merz@vu.nl).

Humans are remarkable in the extent to which they engage in prosocial behavior – and many communities want to encourage their members to help each other and contribute to public goods. According to classical economic theory, an obvious means to do this is by offering incentives¹ (2, 3). Empirical investigations have in fact demonstrated that incentives sometimes encourage prosocial behavior. For example, tax breaks increase charitable giving (4); lottery tickets and other rewards increase blood donations (5, 6). However, incentives sometimes backfire. A range of studies has found incentives to have no effect on the amount of prosocial behavior exhibited; in some cases they have actually reduced prosociality (7–12). For instance, high school students collecting donations took in *less* when they were paid in proportion to the charitable donations they collected than when they were not paid (12).

Indeed, reviews both in economics (13) and in the blood donation literature (1) conclude that the effectiveness of incentives in encouraging prosocial behavior is strongly *context-dependent*. Context effects include: (a) *methodological* effects, e.g., surveys tend to find that people report not being attracted by incentives, whereas experiments and observational studies more often find positive effects of incentives on observed behavior (6); (b) varying effects across incentive *types*, e.g., lottery tickets (5) and time off work (14) encouraged blood donation, but movie tickets (15) did not; and (c) *cross-country* variation in the effectiveness of incentives, e.g., recruitment rates for new blood donors were increased by coupons in the U.S. (16), but remained unchanged when vouchers were offered to Argentinians (17). What explains these inconsistencies, and under what circumstances do incentives encourage prosocial behavior?

Previous theoretical explanations for effects of incentives on prosocial behavior. Richard Titmuss’s classic argument (18) that incentives lead to a deterioration of people’s intrinsic motivation to do good inspired theorizing about the *detrimental* effects of incentives. In particular, it has been suggested that extrinsic rewards cause a decrease in intrinsic motivation, resulting in the phenomenon of “motivational crowding out” (7, 8). Crowding out was said to prevent agents from indulging in altruistic feelings, as a result of introducing incentives (19, 20).

However, these accounts could not explain why incentives sometimes do encourage prosocial behavior, in particular when people are not observed (i.e., in private settings). Thus, another mechanism argued to play a role in determining the effect of incentives on prosocial behavior is *reputational* motivation² (also termed image motivation 21). For instance, participants exerted *more* effort to donate to a charitable organization if they received financial incentives for the amount of effort employed, *if* it was *not visible* to others that they were being paid (21). But when it was *visible* to others, participants actually employed the same or *less* effort than when they were not being paid.

Similar findings emerged from the literature on dictator games (DG), where agents trade off financial reward for themselves (i.e., extrinsic motives) and sharing the endowment with another (intrinsic motivation to act prosocially). The classic paper by Hoffman et al. (22) was one of the first to show that contributions in a DG were much lower in double-blind conditions where neither the other participants nor the experimenters were aware of participants’ choices, compared to when experimenters were able to observe the participants’ behavior. Similarly, participants behaved less prosocially when their behavior had less direct reputational consequences (for instance because there were multiple dictators, 23). In sum, people are more likely to be tempted by an extrinsic reward when their reputation is not at stake.

¹We define an incentive as an “extrinsically motivated reward, which can be monetary or nonmonetary and is used to motivate an action” (1, p.245). Extrinsic rewards are the positive *external* outcomes associated with a given behavior, such as financial rewards. This is opposed to *intrinsic rewards*, i.e., the positive *internal* outcomes of a behavior, such as feeling fulfilled. While *extrinsic* rewards range from financial benefits to status gains, we focus specifically on (a) *financial* rewards (e.g., tax breaks, cash) and (b) non-financial *time* rewards (i.e., time off work).

²We define reputational motivation as the motivation to act derived from the *valuation of reputational outcomes* associated with a behavior, where reputational outcomes are the positive or negative evaluations of a behavior given what is considered *appropriate* in a specific society.

This notion that people care about how they are perceived has also been incorporated into theory. Notably, Andreoni and Bernheim (24) included reputational motivation in their model of dictator game behavior, where agents' behavior is driven by, among other factors, a desire to be perceived as fair. Similarly, Bénabou and Tirole (25) modeled prosocial behavior as being influenced by a desire to be perceived as intrinsically and not extrinsically motivated. These models have been successful in predicting differences between public and private settings, and have offered an explanation for discrepancies between people's self-reported attitudes toward incentives and observed behavior. However, varying effects of different types of incentives and cross-country variation in the effectiveness of incentives remain unexplained.

Social norms govern reputational motivation. Reputation is in fact a core element of explanations for human prosociality in evolutionary biology and social psychology (26–30). In particular, accounts from evolutionary biology posit that reputation formation enables indirect reciprocity, which is critical for the evolution of human cooperation (26); in psychology, reputational concerns are argued to be an important driver for people to engage in individually costly (prosocial) behavior, because positive reputational outcomes (e.g., increased respect) result in long-term benefits (29, 30). Crucially, reputation here is understood to be governed by *social norms*³ (26, 29–32): “[t]he revision of an individual’s reputation depends on the social norms that establish what characterizes a good or bad action [...]” (33, p. 242).

A critical feature of social norms is context dependency. In some contexts *prosocial* norms may dominate, prescribing that resources should be shared with others, e.g., due to societal responsibility (30) or fairness (34). In other contexts, it may be acceptable for people to behave more selfishly (e.g., people do not prefer the 50-50 split in DG when earned wealth is allocated 35). Krupka and Weber (36) provide striking evidence of the subtle context dependency of social norms in variations of the DG. For example, people found a 50-50 split more appropriate when initial allocations to the dictator and recipient were 50-50 than when initial allocations were 100-0, despite payoff equivalence in these two scenarios. In turn, this variation in norms was a good predictor of choices made by unrelated others in these games (36). In addition to this situational context dependency, social norms also depend on *cultural* context (26, 31). For instance, sharing rates in DG vary widely across societies, which in turn has been ascribed to variation in fairness norms (37–39).

Similarly to variation in norms regarding the appropriate behavior in lab experiments, there is variability in norms regarding *real-world prosocial behaviors*. For example, cultural variation exists in need-based communal norms versus conditional exchange norms (40) and in prosocial norms related to environmentalism (41). Moreover, cross-cultural research found significant variation in the *performance* of prosocial behaviors across societies, such as variation in rates of volunteering (42) and charitable giving (43). However, the relationship between *real-world* prosocial behavior and norms (as has been established for lab experiments by 36) remains unclear.

Thus there exists well-documented contextual variation in (prosocial) norms, and we argue that norms applying to the provision of incentives for behaving prosocially are no exception. That is, we deem it plausible that norms vary across incentive types and cultural contexts, and that this has reputational effects which in turn influence behavior. Akin to experimental investigations of the effects of norms on prosocial behavior (36), we aimed to test the relationship between norms (regarding incentives) and *real-world* prosocial behavior. Instead of simply assuming that people want to be perceived as prosocial or as intrinsically motivated as possible (as formalized for example in 25), we propose to explicitly take social norms into account.

The current research. In the following, we detail a theoretical model of prosocial behavior that posits a crucial role of social norms in determining the effects of incentives on prosociality. We empirically test the model on a real-world prosocial behavior: blood

³We define social norms as the informal rules that guide and/or constrain social behaviors, which determine how acceptable a given behavior is viewed by members of a society or group (30).

donation. Ensuring a sufficient supply of blood is very important for all societies, as blood plays an essential part in many medical procedures (e.g., surgeries, cancer treatment) and for producing life-saving drugs. The case of blood donation is also theoretically interesting, as there have been many debates about whether donors should be offered incentives (18), and as a result different countries have adopted various policies regarding incentives for donors (1, 44).

Results

A formal model of prosocial behavior. We implement a formal model of prosocial behavior based on the prominent model of prosocial behavior by Bénabou and Tirole (25) (henceforth BT). We follow BT in positing that prosocial behavior is governed by (a) the direct benefit and (b) the reputational benefit associated with performing the behavior. Specifically, an agent A decides whether to engage in the prosocial behavior B by maximizing his benefits:

$$\text{Max}\{\text{direct benefit} + \text{reputational benefit}\}, B \in \{0, 1\} \quad (1)$$

Here, *direct benefit* is itself composed of three subcomponents: intrinsic motivation $v_a \in [0, 1]$, extrinsic motivation $v_v \in [0, 1]$ and cost $c \in [0, 1]$ (e.g., in terms of time or resources). The latter has a negative effect, whereas the former two have positive effects on motivation to perform B . Importantly, extrinsic motivation v_v can play a role only if an incentive R is provided (i.e., if $R = 1$; $R \in \{0, 1\}$), otherwise $v_v = 0$. *Direct benefit* is thus defined as follows:

$$\text{direct benefit} = B * (v_a + v_v * R) - B * c \quad (2)$$

Reputational benefit represents the benefit (or cost) of being perceived in a certain way by others, i.e., the reputational outcomes related to the behavior. BT theorized *reputational benefit* as depending on VIS , the visibility of the action (i.e., to what extent the behavior is performed in private vs. public), as well as $pref_{v_a}$ and $pref_{v_v}$, the individual’s preference for appearing intrinsically and not extrinsically motivated respectively. Most importantly, reputational motivation depends on the *expected value* of an individual’s intrinsic and extrinsic motivation given their behavior and the extrinsic reward associated with the behavior. Crucially, BT assumed that the expected value of extrinsic motivation $E(v_v)$ is a *cost* whereas the expected value of intrinsic motivation $E(v_a)$ is a *benefit*.

Notably, expected intrinsic motivation $E(v_a|R, B)$ as formulated in the model has a positive effect irrespective of whether incentives are offered. Individuals who engage in the prosocial activity have a higher expected intrinsic motivation than individuals who do not behave prosocially. However, the same does not hold for the cost associated with higher expected extrinsic motivation. In the presence of incentives, individuals engaging in the prosocial activity are perceived to have higher extrinsic motivation than individuals not engaging in the prosocial activity. The model thus demonstrates that ”motivational crowding-out” can arise endogenously, as the cost of being perceived as extrinsically motivated⁴ offsets the benefits of extrinsic motivation itself.

In the following, we deviate from BT’s model by introducing two additional parameters capturing social norms. We replace the hard-coded cost (benefit) to being perceived as extrinsically (intrinsically) motivated with two social norm variables in order to account for contextual variation in how acceptable members of society view appearing intrinsically or extrinsically motivated by a given incentive offered for a particular prosocial act. We call this social norm variable S and define it as the *sum of reputational costs and benefits as dictated by the prevalent social norms associated with a given behavior*. This applies to both expected intrinsic motivation (i.e., $S_{v_a} \in [-1, 1]$)

⁴In addition to a reputational cost from appearing extrinsically motivated, decreased perceived *intrinsic* motivation as a result of receiving an incentive presents a further, albeit minor, reputational cost. This is because $E(v_a|R = 0, B = 1)$ is slightly larger than $E(v_a|R = 1, B = 1)$; the extent of this reputational cost depends on the costliness of performing the prosocial behavior (see SI Appendix Fig. S2).

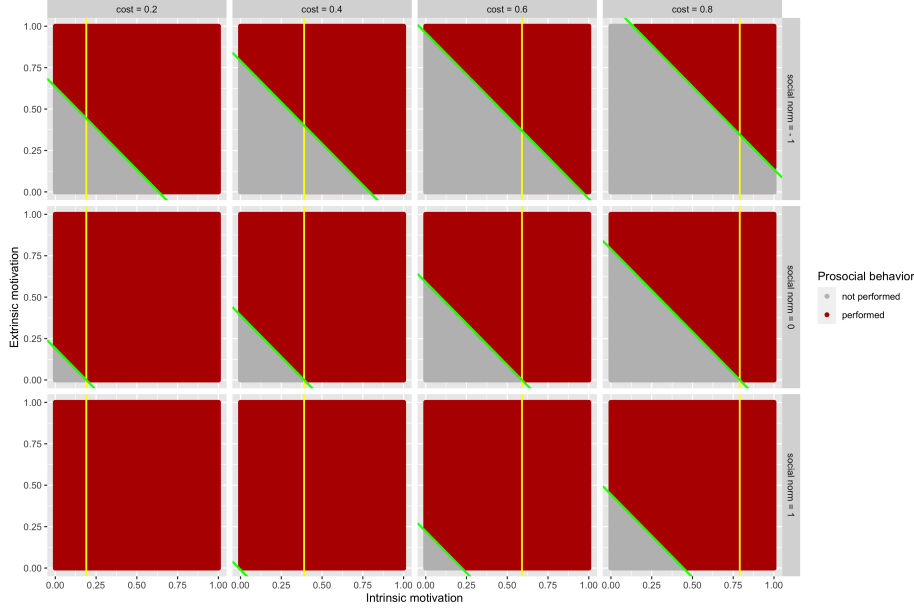


Figure 1: Simulated behavior of individuals differing in intrinsic (x-axis) and extrinsic (y-axis) motivation for different values of social norms (rows) and cost (columns) when incentives are offered. In the presence of incentives, the value of the social norms variable S_{v_v} determines the shift on the x-axis of the diagonal line dividing those behaving prosocially and those not (green line; see description in main text). Parameter settings: $VIS = pref_{v_a} = pref_{v_v} = 1$; $S_{v_a} = 0$; $v_a \sim unif\{0, 1\}$ and $v_v \sim unif\{0, 1\}$.

and expected extrinsic motivation (i.e., $S_{v_v} \in [-1, 1]$). From this follows our definition of reputational benefit:

$$reputational\ benefit = VIS[S_{v_a} * pref_{v_a} * E(v_a|R, B) + S_{v_v} * pref_{v_v} * E(v_v|R, B)] \quad (3)$$

To illustrate the effect of our model modification, let us consider the effect of varying levels of S_{v_v} on prosocial behavior in the presence of incentives. When S_{v_v} takes on the value 0, as depicted in the second row of Fig. 1, a prosocial act is performed if the additive effects of intrinsic and extrinsic motivation outweigh the costs. That is, individuals whose *intrinsic* motivation is larger than the costs of displaying the behavior – who are located to the right of the yellow line – perform the prosocial act. And in addition, individuals with insufficient intrinsic, but some *extrinsic* motivation – who are located between the yellow and green diagonal lines – are also motivated to display the behavior. However, when S_{v_v} takes on a *negative* value (Fig. 1, first row), the diagonal threshold dividing those behaving prosocially and those not (green line) shifts toward the right on the x-axis. This is due to agents incurring a reputational cost arising from social norms that characterize it as “bad” that the given prosocial act is performed while the agents receive a specific incentive⁵. On the other hand, when S_{v_v} is positive (Fig. 1, third row), the green threshold shifts to the left on the x-axis, reflecting high levels of prosocial behavior due to the combined positive effect of intrinsic, extrinsic and reputational motivation.

⁵Note that we take *social norms regarding incentive X* as a proxy for our theoretical construct *social norms for being extrinsically motivated by incentive X* (S_{v_v}), because the main reputational consequence of introducing incentives is that individuals appear more extrinsically motivated (see SI Appendix Fig. S2).

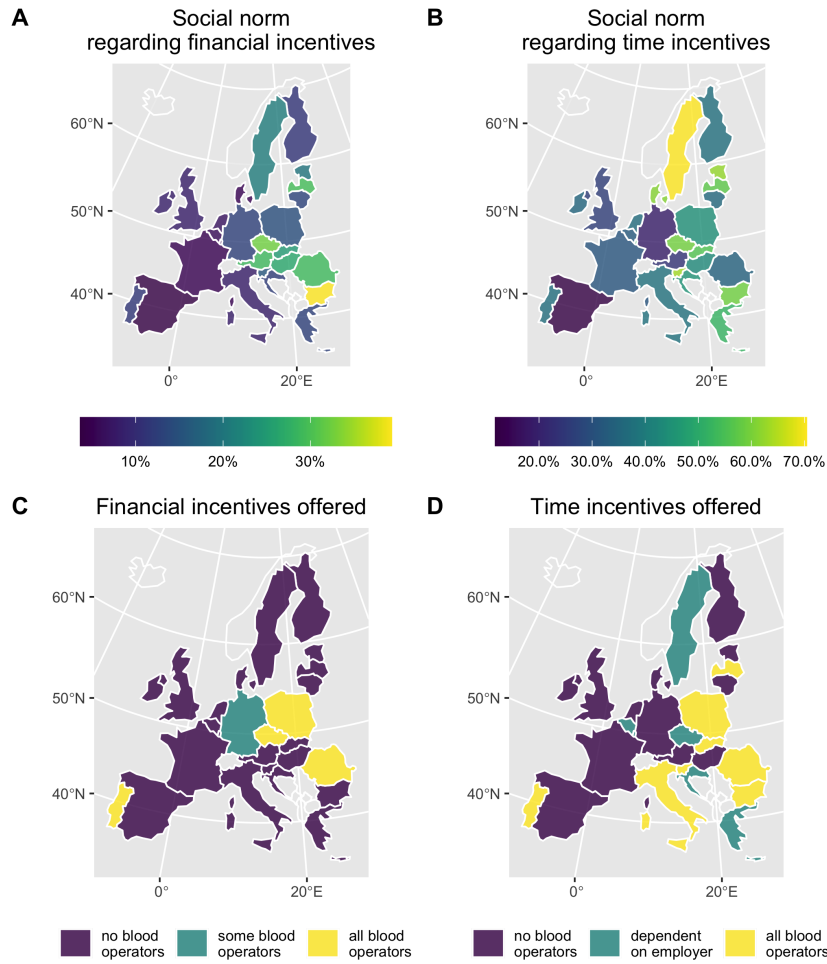


Figure 2: Maps displaying blood donor incentives and norms regarding the acceptability of incentives across Europe. **(A)** Social norm regarding the acceptability of *financial* incentives for blood donation. **(B)** Social norm regarding the acceptability of *time* incentives for blood donation. **(C)** Financial incentives and **(D)** time incentives offered to blood donors (where the incentive may be provided by *all*, only *some* or *none* of the blood operators in a given country; in the case of *time* incentives, the incentive may be offered only to *donors whose employer allows it*).

Testing the model on a real-world prosocial behavior. We test our formal model empirically on the real-world prosocial behavior of blood donation⁶. To this end we employed large-scale survey data from the Eurobarometer (45), which was obtained from representative samples of 28 European countries ($n = 26\,192$). Overall, 38.4% ($n = 10195$) of all respondents included had ever donated blood. This varied substantially across countries (see SI Appendix Fig. S3). Similarly, we found strong variation in social norms regarding incentives (see Fig. 2A and B; social norms regarding financial incentives: $X^2(27) = 1970.7, p < 0.001$; social norms regarding time incentives: $X^2(27) = 2436.6, p < 0.001$). However, social norms regarding *time* incentives were more positive (mean = 0.43) than those regarding *financial* incentives (mean = 0.15; Mann–Whitney $U(n_1 = n_2 = 28) = 736, p < 0.001$). Moreover, both financial and time incentives offered to blood donors varied across countries (see 2C and D). See SI Appendix for more descriptive information and figures.

We found a positive main effect of social norms regarding *financial* incentives on blood donation propensity ($b = 0.125, p < 0.05$), but a negative main effect of financial incentives themselves ($b = -0.512, p < 0.01$). As illustrated in Fig. 3A, these results indicate that respondents from countries where incentives are offered are, overall, less likely to have donated blood. Yet, the likelihood of having donated was higher in countries with more positive norms regarding financial incentives. We did not find a significant interaction effect between financial incentives and social norms. However, we did find a positive interaction between *time* incentives and social norms ($b = 0.484, p < 0.01$). That is, people living in a country providing time incentives are more likely to donate blood when social norms regarding time incentives are more positive (see 3B).

In addition, the full model revealed a positive effect of self-reported intrinsic motivation ($b = 1.346, p < 0.0001$), and a negative effect of cost of donating (in terms of living in rural areas) on giving blood ($b = -0.091, p < 0.05$). Individuals who reported

⁶Specifically, we are interested in examining how prosocial behavior is related to incentives and social norms, as well as to intrinsic and extrinsic motivation. Thus, our focus is on the model parameters v_a, v_v, c, R and S_{v_v} . The effects of other model parameters on prosociality are documented elsewhere (e.g., visibility 21).

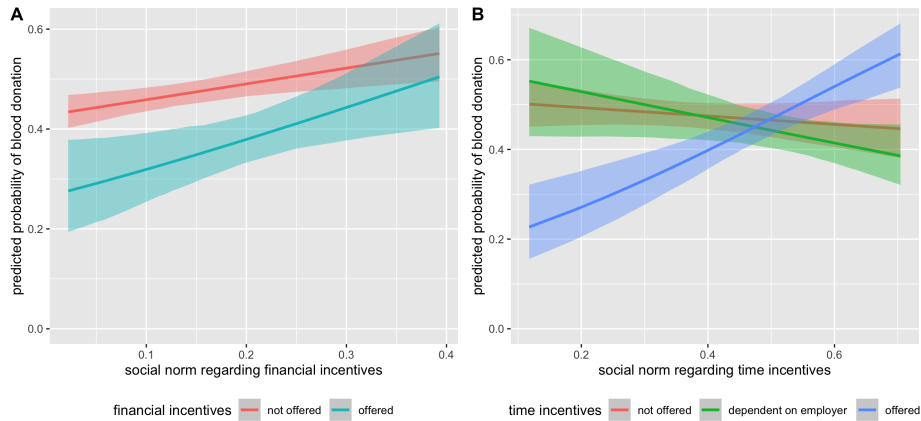


Figure 3: Predictive margins of donating blood for a representative European woman (i.e., mean-aged, living with a partner in a small town, employed, having completed secondary-level education, mean number of children, rarely having difficulty paying bills, intrinsically but not extrinsically motivated) as a function of (A) social norms regarding *financial* incentives, split by financial incentives (where incentives are offered to all blood donors in a given country or not at all) and (B) social norms regarding *time* incentives, split by time incentives (where incentives are offered to all blood donors in a given country, only to donors whose employer allows it, or not at all). Prediction bands indicate 80% prediction intervals.

being motivated by the desire to help those in need, alleviate shortages or support medical research and those living in urban areas are more likely to have donated blood⁷ (see SI Appendix Tab. S3 and S4 for full model results). The effect of self-reported extrinsic motivation (in terms of getting something in return as a motive to donate) was in the expected direction but non-significant ($b = 0.110, p = 0.078$).

Discussion

In this paper we extended a model of prosocial behavior that posits an additive effect of intrinsic motivation, extrinsic motivation, reputational motivation, and costs. Our key theoretical contribution is the introduction of social norms. We argue that reputational costs or benefits are a product of social norms, i.e., how acceptable certain behaviors are viewed by members of society. In our view this theoretical modification allows us to account for the inconsistent effects of incentives on prosocial behavior across contexts. We tested the model on a real-world prosocial behavior, namely blood donation, and found a positive association between country-level social norms for financial incentives and blood donation: the more positive the social norm, the higher the probability of having donated blood. While financial incentives were negatively associated with blood donation, time incentives were on their own unrelated to the likelihood of donating. Instead, the effect of time incentives depended on social norms; i.e., in countries where time incentives are offered, individuals' likelihood of donating was higher when social norms regarding these incentives were positive than when social norms were more negative.

Akin to previous studies (see 1, 13), we found that incentives have inconsistent effects on prosocial behavior. Whereas time incentives alone were unrelated to an individual's likelihood of having donated blood, financial incentives were negatively associated with the propensity to donate blood. However, the differences in the relationship between financial and non-financial incentives and prosociality can in fact be explained by norms – i.e., the social norm regarding the incentive predicts whether the probability of donating increases. On the one hand, time incentives were viewed rather negatively in some of the countries offering them, and respondents from these countries were less likely to donate compared to respondents in countries without time incentives. On the other hand, respondents in countries where social norms regarding time incentives were more positive were *more* likely to donate than comparable respondents from countries not offering time incentives. Similarly for financial incentives, the likelihood of donating increased if social norms were more positive, even though financial incentives themselves had a negative effect on the likelihood of donating blood. As such, our findings highlight the relevance of BT's theory in a large population and for a real-world prosocial behavior. But more importantly, our conclusions suggest that our theoretical modification derived from interdisciplinary cross-pollination – the integration of social norms – provides a fruitful addition to our understanding of the relationship between incentives and prosocial behavior.

While we found that more positive norms regarding *time* incentives are associated with a higher predicted probability of donating *if* time incentives are offered, our observation that social norms regarding *financial* incentives are positively associated with donation behavior *irrespective of whether incentives are offered* is unexpected. A possible explanation for this may be that more positive norms regarding financial incentives capture additional societal attitudes that impact donation behavior, e.g., valuation of the donor status generally.

The most direct implication that follows from our findings is that incentives *can* be associated with increased levels of prosocial behavior – *if* social norms regarding the incentive are sufficiently positive. Although clever experimental interventions have been designed to avert the reputational *costs* arising from offering incentives (e.g., 46), our study suggests that there need not be such *costs* for all combinations of

⁷As a robustness check, we implemented the same model on a subset of respondents, excluding those who indicated they are not willing to donate in the future (these respondents were not asked about motivational factors; see SI Appendix). For this sample we also found a positive, albeit smaller, effect of intrinsic motivation on blood donation ($b = 0.091, p < 0.05$).

incentives and prosocial activities, but that reputational consequences depend on the relevant norms. Thus, experimenting with different types of incentives (which may be perceived as more or less appropriate in a given society) seems promising as a way to find incentives that encourage societally desirable behavior with less or no reputational cost. In addition, it may be possible to increase the effectiveness of incentives by changing how incentives are framed. For instance, blood operators might be able to promote views emphasizing how fair it is that donors receive time off work due to their critical role in the medical system.

While our findings entail important theoretical and policy implications, open questions remain for future research to explore. Even though we found that a relatively simple measure of social norms explains interesting empirical patterns of real-world prosocial behavior, a more fine-grained measure of social norms (e.g., according to 36, 47) would offer further important insights. For instance, a measure targeting more specifically the acceptability of *being extrinsically motivated* by incentive X to perform behavior Y could reveal precisely which norms govern the reputational consequences of receiving an incentive for behaving prosocially. In addition, the operationalization of intrinsic and extrinsic motivation could be improved, as the Eurobarometer employed a selection criterion for these measures (see SI Appendix). Although intrinsic and extrinsic motivation were not our main concern in this paper, further examination of the interplay of various motivational factors and social norms seems promising. Lastly, we cannot draw any causal conclusions from our observational methods. Our finding of lower donation rates in countries offering financial incentives does not necessarily mean that incentives do not encourage donation, and the association between country-level social norms and donation rates may be due to other factors. Experimental investigations could further our understanding of the effects of norms on prosocial behavior in the presence of incentives, including to what extent norms can be manipulated in the lab and field.

Nevertheless, our unique approach allowed us to investigate the effect of social norms “in the wild”; That is, we were able to examine empirically elicited norms across countries which are typically very persistent and have a long cultural evolutionary history. Our observation that these norms offer an explanation both for cross-country variation and differences across types of incentives highlights not only the theoretical merits of incorporating social norms, but also offers concrete implications for policy-makers regarding how to encourage and influence prosocial behavior more generally.

Materials and Methods

All data and code for running the formal model and analyses, as well as our preregistration⁸, are available at <https://osf.io/bf2kt/>.

Participants and procedure. We employed comparative data from the 2014 wave of the Eurobarometer (45), which is a repeated cross-sectional survey conducted on representative samples from European Union member states. In 2014, the survey was performed in 28 countries. A multi-stage random (probability) sampling design was employed and approximately 1,000 face-to-face interviews were completed in each of the countries. Respondents were residents in the respective country, had sufficient command of the national language(s), and were 15 years or older. Study protocols were approved by the European Commission and all participants provided informed consent. A detailed description of all variables, including the wording of survey items and measurement scales, is presented in SI Appendix, Tab. S5.

⁸Note that only analyses pertaining to donation of *blood* (not other substances of human origin) are presented here. We deviate from the preregistration in two respects. Firstly, instead of creating a “conditional” social norm and extrinsic motivation variable (i.e., setting these variables to 0 if no incentive is offered), we employ interaction terms between incentives and social norms / extrinsic motivation. Secondly, we did not run a robustness check excluding those countries not offering incentives, because this would have resulted in too small a sample size at the country-level.

Blood donation. The dependent variable in our analyses is blood donation during one’s lifetime (yes/no). Participants in the Eurobarometer were asked whether they have ever donated substances of human origin, including blood. All categorical variables, including the dependent variable, are dummy-coded.

Incentives. Based on responses to the expert survey (see below), we constructed two country-level incentive variables: (a) financial incentives (i.e., cash or other high-value incentives, for example tax deduction) and (b) time incentives (i.e., time off work). There were three categories of financial incentives: (1) offered by *no* blood operators, (2) *some* blood operators, and (3) all blood operators. Time incentives also had three levels: (1) offered by *no* blood operators, (2) offered *dependent on the employer* (i.e., in some countries employers could grant donors permissions to donate during working hours, but this did not hold for all donors), and (3) offered by all blood operators.

Social norms regarding incentives. As social norms by definition apply to the (societal) group level, we operationalized *country-level* acceptability ratings of incentives rather than using respondents’ individual attitudes. Specifically, we aggregated the responses of individual respondents regarding the acceptability of receiving incentives for blood donation in each of the 28 European countries. This resulted in a country-level variable which ranges from 0 (i.e., 0% of respondents from a given country find financial/time incentives acceptable) to 1 (i.e., 100% of respondents from a given country find financial/time incentives acceptable). We computed social norms separately for financial and time incentives.

Intrinsic and extrinsic motivation. Intrinsic motivation is operationalized as the mentioning (yes/no) of an intrinsic motivational factor as a reason for donating blood. Intrinsic motivation is coded “yes” if one of three intrinsic motivational factors are mentioned: (a) to help other people in need, (b) to alleviate shortages of blood, or (c) to support medical research. Extrinsic motivation both for financial and time incentives is operationalized as the mentioning (yes/no) of an extrinsic factor as a reason for donating or being willing to donate blood (namely “to receive something in return for you or your relatives”). We additionally included two other measures of extrinsic motivation for *financial* incentives: (1) subjective difficulty in paying bills (three levels: almost never/never, from time to time, most of the time; we assumed higher extrinsic motivation would follow from increased difficulty in paying bills) and (2) lack of current employment status (not employed/employed; we assumed higher extrinsic motivation for *financial* incentives would follow from being *unemployed*). For time incentives, we included the presence of employment as an extrinsic motivator (employed/not employed; we assumed extrinsic motivation for *time* incentives would be higher for those employed).

Additional covariates. We controlled for socio-demographic factors, which previous studies had shown to be related to propensity to donate (see (48) for a recent review on donor demographics). We included (a) age (as our dependent variable is dependent on years lived), (b) gender (male/female), (c) education (in years, categorized in four levels), and (d) partner status (cohabitating: yes/no). In addition, we controlled for two factors which represent costs of donating: (e) time constraints due to childcare responsibilities (operationalized as number of children in household), and (f) travel costs due to living in rural areas (i.e., rural areas have fewer blood collection locations; three levels: large town; small/middle-sized town; rural area).

Expert survey. We conducted a survey of incentive policies among specialists at national blood collection agencies across Europe. Experts reported on the type and value of incentives provided in their country by means of a telephone interview or in writing. We explicitly asked experts not to mention low-value incentives (as it is common for donors to receive small gifts such as T-shirts or mugs); we are interested only in high-value (non-symbolic) incentives. Incentive data for the Czech Republic,

France, and the Netherlands were taken from a recently published study on donor incentives across the globe (49).

Analytical strategy. Respondents younger than 18 years were excluded since most EU countries do not permit people under 18 to donate blood⁹. We employed logistic multilevel mixed models for all main analyses, where individuals were nested in countries. The models included individual-level variables (e.g., intrinsic motivation) and country-level variables (e.g., incentives) as fixed effects, as well as allowing the intercept to vary by country. The estimation technique is maximum likelihood. Missing values were handled by list-wise deletion. We ran separate models for financial and time incentives. The predictors of interest are intrinsic motivation, cost, extrinsic motivation in interaction with financial/time incentives, and social norms regarding financial/time incentives in interaction with financial/time incentives. Prediction intervals are based on fixed coefficients and observation-level errors. All analyses were run with R version 4.0.2 using the packages lme4 (modeling), merTools (predictions), tidyverse (data transformation; plotting) and rnatuarearth (maps) (50–53).

References

- [1] K. Chell, T. E. Davison, B. Masser, and K. Jensen, “A systematic review of incentives in blood donation,” *Transfusion*, vol. 58, no. 1, pp. 242–254, 2018. Publisher: Wiley Online Library.
- [2] J. S. Mill, “Principles Of Political Economy,” *Proceedings of the American Academy of Science*, vol. 1881, p. 417, 1882.
- [3] A. Marshall, *Principles of Economics, by Alfred Marshall*. Macmillan and Company, 1890.
- [4] N. J. Duquette, “Do tax incentives affect charitable contributions? Evidence from public charities’ reported revenues,” *Journal of Public Economics*, vol. 137, pp. 51–69, 2016. Publisher: Elsevier.
- [5] L. Goette and A. Stutzer, “Blood donations and incentives: Evidence from a field experiment,” *Journal of Economic Behavior & Organization*, vol. 170, pp. 52–74, 2020. Publisher: Elsevier.
- [6] N. Lacetera, M. Macis, and R. Slonim, “Economic rewards to motivate blood donations,” *Science*, vol. 340, no. 6135, pp. 927–928, 2013. Publisher: American Association for the Advancement of Science.
- [7] E. L. Deci, “Effects of externally mediated rewards on intrinsic motivation.,” *Journal of Personality and Social Psychology*, vol. 18, no. 1, p. 105, 1971. Publisher: American Psychological Association.
- [8] E. L. Deci, R. Koestner, and R. M. Ryan, “A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation.,” *Psychological bulletin*, vol. 125, no. 6, p. 627, 1999.
- [9] C. Mellström and M. Johannesson, “Crowding out in blood donation: was Titmuss right?,” *Journal of the European Economic Association*, vol. 6, no. 4, pp. 845–863, 2008. Publisher: Oxford University Press.
- [10] J. Heyman and D. Ariely, “Effort for payment: A tale of two markets,” *Psychological Science*, vol. 15, no. 11, pp. 787–793, 2004. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

⁹In addition, respondents from Germany were excluded from analyses pertaining to financial incentives since Germany was the only country where blood operators across the country differed in the financial incentives provided. The results are quantitatively equivalent when Germany is included and regrouped either as offering, or not offering, financial incentives.

- [11] U. Gneezy and A. Rustichini, “A fine is a price,” *The Journal of Legal Studies*, vol. 29, no. 1, pp. 1–17, 2000. Publisher: The University of Chicago Press.
- [12] U. Gneezy and A. Rustichini, “Pay enough or don’t pay at all,” *The Quarterly Journal of Economics*, vol. 115, no. 3, pp. 791–810, 2000. Publisher: MIT Press.
- [13] U. Gneezy, S. Meier, and P. Rey-Biel, “When and why incentives (don’t) work to modify behavior,” *Journal of Economic Perspectives*, vol. 25, no. 4, pp. 191–210, 2011.
- [14] N. Lacetera and M. Macis, “Time for blood: the effect of paid leave legislation on altruistic behavior,” *The Journal of Law, Economics, & Organization*, vol. 29, no. 6, pp. 1384–1420, 2013.
- [15] D. Royse, “Exploring ways to retain first-time volunteer blood donors,” *Research on Social Work Practice*, vol. 9, no. 1, pp. 76–85, 1999.
- [16] J. R. Ferrari, R. C. Barone, L. A. Jason, and T. Rose, “The use of incentives to increase blood donations,” *The Journal of social psychology*, vol. 125, no. 6, pp. 791–793, 1985.
- [17] V. Iajya, N. Lacetera, M. Macis, and R. Slonim, “The effects of information, social and financial incentives on voluntary undirected blood donations: Evidence from a field experiment in argentina,” *Social Science & Medicine*, vol. 98, pp. 214–223, 2013.
- [18] R. M. Titmuss, “The Gift Relationship: From Human Blood to Social Policy (New York),” *Pantheon*, vol. 74, 1971.
- [19] B. S. Frey and F. Oberholzer-Gee, “The cost of price incentives: An empirical analysis of motivation crowding-out,” *The American Economic Review*, vol. 87, no. 4, pp. 746–755, 1997. Publisher: JSTOR.
- [20] B. S. Frey and R. Jegen, “Motivation crowding theory,” *Journal of Economic Surveys*, vol. 15, no. 5, pp. 589–611, 2001. Publisher: Wiley Online Library.
- [21] D. Ariely, A. Bracha, and S. Meier, “Doing good or doing well? Image motivation and monetary incentives in behaving prosocially,” *American Economic Review*, vol. 99, no. 1, pp. 544–555, 2009.
- [22] E. Hoffman, K. McCabe, and V. L. Smith, “Social distance and other-regarding behavior in dictator games,” *The American Economic Review*, vol. 86, no. 3, pp. 653–660, 1996. Publisher: JSTOR.
- [23] J. Dana, R. A. Weber, and J. X. Kuang, “Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness,” *Economic Theory*, vol. 33, no. 1, pp. 67–80, 2007. Publisher: Springer.
- [24] J. Andreoni and B. D. Bernheim, “Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects,” *Econometrica*, vol. 77, no. 5, pp. 1607–1636, 2009. Publisher: Wiley Online Library.
- [25] R. Bénabou and J. Tirole, “Incentives and prosocial behavior,” *American Economic Review*, vol. 96, no. 5, pp. 1652–1678, 2006.
- [26] M. A. Nowak and K. Sigmund, “Evolution of indirect reciprocity,” *Nature*, vol. 437, no. 7063, pp. 1291–1298, 2005. Publisher: Nature Publishing Group.
- [27] M. A. Nowak and K. Sigmund, “Evolution of indirect reciprocity by image scoring,” *Nature*, vol. 393, no. 6685, pp. 573–577, 1998. Publisher: Nature Publishing Group.
- [28] E. Fehr and U. Fischbacher, “Social norms and human cooperation,” *Trends in cognitive sciences*, vol. 8, no. 4, pp. 185–190, 2004.

- [29] A. Morris and F. Cushman, “A common framework for theories of norm compliance,” *Social Philosophy & Policy*, vol. 35, no. 1, pp. 101–127, 2018.
- [30] R. B. Cialdini and M. R. Trost, “Social influence: Social norms, conformity and compliance,” 1998. Publisher: McGraw-Hill.
- [31] R. D. Alexander, *The biology of moral systems*. Transaction Publishers, 1987.
- [32] M. Sherif, “The psychology of social norms.,” 1936. Publisher: Harper.
- [33] F. P. Santos, F. C. Santos, and J. M. Pacheco, “Social norm complexity and past reputations in the evolution of cooperation,” *Nature*, vol. 555, no. 7695, pp. 242–245, 2018. Publisher: Nature Publishing Group.
- [34] E. Fehr and K. M. Schmidt, “A theory of fairness, competition, and cooperation,” *The Quarterly Journal of Economics*, vol. 114, no. 3, pp. 817–868, 1999. Publisher: MIT Press.
- [35] T. L. Cherry, P. Frykblom, and J. F. Shogren, “Hardnose the dictator,” *American Economic Review*, vol. 92, no. 4, pp. 1218–1221, 2002.
- [36] E. L. Krupka and R. A. Weber, “Identifying social norms using coordination games: Why does dictator game sharing vary?,” *Journal of the European Economic Association*, vol. 11, no. 3, pp. 495–524, 2013.
- [37] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath, “In search of homo economicus: behavioral experiments in 15 small-scale societies,” *American Economic Review*, vol. 91, no. 2, pp. 73–78, 2001.
- [38] J. Henrich, S. J. Heine, and A. Norenzayan, “The weirdest people in the world?,” *Behavioral and Brain Sciences*, vol. 33, no. 2-3, pp. 61–83, 2010. Publisher: Cambridge University Press.
- [39] Y. Kawamura and T. Kusumi, “Altruism does not always lead to a good reputation: A normative explanation,” *Journal of Experimental Social Psychology*, vol. 90, p. 104021, 2020. Publisher: Elsevier.
- [40] J. G. Miller, H. Akiyama, and S. Kapadia, “Cultural variation in communal versus exchange norms: Implications for social support.,” *Journal of Personality and Social Psychology*, vol. 113, no. 1, p. 81, 2017. Publisher: American Psychological Association.
- [41] B. Torgler, B. S. Frey, and C. Wilson, “Environmental and pro-social norms: evidence on littering,” *The BE Journal of Economic Analysis & Policy*, vol. 9, no. 1, 2009. Publisher: De Gruyter.
- [42] S. Ruiter and N. D. De Graaf, “National context, religiosity, and volunteering: Results from 53 countries,” *American Sociological Review*, vol. 71, no. 2, pp. 191–210, 2006. Publisher: Sage Publications Sage CA: Los Angeles, CA.
- [43] T. K. Bauer, J. Bredtmann, and C. M. Schmidt, “Time vs. money—the supply of voluntary labor and charitable donations across europe,” *European Journal of Political Economy*, vol. 32, pp. 80–94, 2013.
- [44] K. Healy, “Embedded altruism: blood collection regimes and the European Union’s donor population,” *American Journal of Sociology*, vol. 105, no. 6, pp. 1633–1657, 2000. Publisher: The University of Chicago Press.
- [45] “European Commission, Data from Eurobarometer 82.2. GESIS Data Archive for the Social Sciences. ZA5931 Datafile Version 3.0.0. Available at doi.org/10.4232/1.12999,” 2018.
- [46] E. L. Kirgios, E. H. Chang, E. E. Levine, K. L. Milkman, and J. B. Kessler, “Forgoing earned incentives to signal pure motives,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 29, pp. 16891–16897, 2020.

- [47] H. Rauhut and F. Winter, “A sociological perspective on measuring social norms by means of strategy method experiments,” *Social Science Research*, vol. 39, no. 6, pp. 1181–1194, 2010. Publisher: Elsevier.
- [48] T. W. Piersma, R. Bekkers, E. F. Klinckenberg, W. L. De Kort, and E.-M. Merz, “Individual, contextual and network characteristics of blood donors and non-donors: a systematic review of recent literature,” *Blood Transfusion*, vol. 15, no. 5, p. 382, 2017. Publisher: SIMTI Servizi.
- [49] M. P. Zeller, D. Ellingham, D. Devine, M. Lozano, P. Lewis, E. Zhiburt, L. van der Linde, M. Goldman, A. Nakamura, S. Inoue, and others, “Vox Sanguinis International Forum on Donor Incentives.,” *Vox Sanguinis*, 2020.
- [50] D. Bates, D. Sarkar, M. D. Bates, and L. Matrix, “The lme4 package,” *R package version*, vol. 2, no. 1, p. 74, 2007.
- [51] J. E. Knowles, C. Frederick, and M. J. E. Knowles, “Package ‘mertools’,” 2016.
- [52] H. Wickham, M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Golemund, A. Hayes, L. Henry, J. Hester, *et al.*, “Welcome to the tidyverse,” *Journal of Open Source Software*, vol. 4, no. 43, p. 1686, 2019.
- [53] A. South, “Rnaturalearth: world map data from natural earth,” *R package version 0.1. 0*, 2017.

Appendix

Supplementary results

Theoretical model: Direct benefit

As described in the main text, we follow Bénabou and Tirole (25, henceforth BT) and assume that the *direct benefit* gained from performing a prosocial behavior B is dependent on an individual's intrinsic motivation v_a , extrinsic motivation v_v , as well as the costs c and incentive R associated with performing the behavior. We defined direct benefit as follows:

$$\text{direct benefit} = B * (v_a + v_v * R) - B * c \quad (4)$$

In the absence of extrinsic rewards (i.e., $R = 0$), only intrinsic motivation v_a and personal cost c determine whether the prosocial action is performed. Personal cost here refers to the individual-level cost associated with the activity, for instance in terms of time, money, and resources. Although costliness of the behavior varies across individuals, it is also conceivable that some activities are highly costly for many people (such as donating a kidney) and are thus performed only by very few who have exceptionally high intrinsic motivation and relatively low cost. On the other hand, some behaviors may be associated with relatively low costs for most people (such as helping someone who tripped and fell) and are therefore performed by most people who have sufficiently high intrinsic motivation. Assuming that the prosocial behavior is performed if *direct benefit* is greater for $B = 1$ than for $B = 0$, the effect of cost¹⁰ is illustrated in Fig. S 1: The first row depicts the effect of varying the cost of the prosocial activity in the absence of incentives. In the upper leftmost facet, most people engage in the behavior (red), because relatively low levels of intrinsic motivation suffice to engage in the relatively uncostly behavior. On the upper rightmost facet is the other extreme, where only the most highly intrinsically motivated individuals perform the prosocial action given the high associated costs.

When incentives are offered (i.e., $R = 1$), the decision to perform the prosocial behavior is governed both by extrinsic and intrinsic motivation as well as costs. This results in a higher total number of individuals that join in performing the behavior, because some individuals are motivated by the extrinsic reward. The bottom row in Fig. S 1 illustrates this effect: the dividing line between those who engage in the prosocial behavior and those who do not virtually tips over toward the left side, as individuals with insufficient intrinsic motivation but some extrinsic motivation now choose to perform the behavior as well.

Theoretical model: Reputational benefit

BT (25) further postulate that *reputational motivation* influences the decision to engage in a prosocial activity or not. As described in the main text, reputational motivation is dependent on an agent's expected extrinsic motivation $E(v_v)$ and expected intrinsic motivation $E(v_a)$ given the rewards offered (where the former represents a cost and the latter represents a benefit for the agent's reputation), as well as $pref_{v_a}$ and $pref_{v_v}$ (i.e., the strength of preference to appear intrinsically motivated and not extrinsically motivated) and VIS (the visibility of the action). Reputational motivation according to BT (25) is then defined as follows:

$$\text{reputational benefit} = VIS[pref_{v_a} * E(v_a|R, B) - pref_{v_v} * E(v_v|R, B)] \quad (5)$$

As explained in the main text, the first component of *reputational benefit* (namely $VIS * pref_{v_a} * E(v_a|R, B)$); of which the driving factor is the expected value of intrinsic motivation $E(v_a|R, B)$) always has a positive effect: Individuals who engage in the prosocial activity have a higher expected intrinsic motivation than individuals who do not behave prosocially, irrespective of whether an extrinsic reward is offered. That is,

$$E(v_a|R = 0, B = 1) > E(v_a|R = 0, B = 0) \quad (6)$$

and

$$E(v_a|R = 1, B = 1) > E(v_a|R = 1, B = 0) \quad (7)$$

¹⁰For illustrative purposes, cost is here assumed to be the same for all individuals. This is however not a constraint in our model, where cost is able to vary across people.

Yet, the same does not hold for the cost associated with higher expected extrinsic motivation: In the presence of incentives, the expected value for extrinsic motivation is in fact higher for individuals engaging in the prosocial activity than not, and this is not the case when there are no incentives provided:

$$E(v_v|R = 0, B = 1) = E(v_v|R = 0, B = 0) \quad (8)$$

and

$$E(v_v|R = 1, B = 1) > E(v_v|R = 1, B = 0) \quad (9)$$

Increased levels of expected extrinsic motivation when incentives are present, combined with ascribing a cost to being perceived as extrinsically motivated, thus introduces a key negative reputational consequence resulting from offering incentives for a prosocial behavior. In addition to this reputational cost from increased *expected extrinsic motivation*, there is a minor reputational cost resulting from decreased *expected intrinsic motivation* associated with performing a prosocial behavior for which an incentive is provided, compared to performing a prosocial behavior for which no incentive is offered. That is, the expected intrinsic motivation of an agent performing a prosocial act is always higher than that of another agent not performing the behavior, but the *extent* of this reputational benefit depends on whether an incentive is offered: Expected intrinsic motivation when performing a prosocial act for which an incentive is offered (i.e., $E(IM|R = 1, B = 1)$) is lower than the expected intrinsic motivation when performing a prosocial act for which *no* incentive is offered (i.e., $E(IM|R = 0, B = 1)$). The difference in expected intrinsic motivation for these two scenarios is dependent on costliness of performing the behavior (see Fig. S 2, blue line), such that the relative reputational cost from expected intrinsic motivation is higher the lower the costliness of the prosocial behavior. Similarly, the size of the reputational cost due to expected extrinsic motivation is also dependent on costliness of the prosocial behavior (see Fig. S 2, red line). While the reputational costs induced by incentives in terms of expected *intrinsic* motivation increase as costliness of performing the prosocial act increases, the costs induced by incentives in terms of expected *extrinsic* motivation decrease with increasing costliness of performing the prosocial act.

In sum, the assumptions by BT (25) that (a) reputational *costs* are associated with being perceived as extrinsically motivated and (b) reputational *benefits* are associated with being perceived as intrinsically motivated, result in reputational *costs* for agents who perform a prosocial behavior in the presence of incentives. These reputational costs induced by incentives are largely driven by costs in terms of being perceived as more extrinsically motivated, but there is also a minor cost due to being perceived as less intrinsically motivated.

Empirical results: Descriptives

Donation rates ranged from 22.8% of respondents from Portugal having ever donated blood to 52.9% of respondents from France having donated blood (see Fig. S 3). Table S 1 displays descriptives of demographic variables and predictors of interest. The correlation matrix of predictors of interest is visualized in S 5. Figure S 4 further displays the geographic distribution of country-level mean levels of intrinsic and extrinsic motivation, which also showed substantial variation (intrinsic motivation: $X^2(27) = 1501.5, p < 0.001$; extrinsic motivation: $X^2(27) = 408.58, p < 0.001$). Notably, levels of intrinsic motivation to donate (see S 4A) were high across all countries (country-level means ranging from 41.3% in Bulgaria to 88.0% in Sweden), albeit higher in Western and Northern European countries compared to Eastern European countries. In contrast, mentioning the extrinsic factor “getting something in return” as motivation to donate was relatively rare (see S 4B; country-level means range from 1.3% in Italy to 12.7% in Finland).

Incentives The expert survey revealed that only few European countries are offering high-value *financial* incentives to blood donors: In Bulgaria, the Czech Republic, Poland, and Portugal all blood operators offer high-value incentives to donors. In Germany, some blood operators offer incentives (e.g., hospital-based blood operators), whereas others (e.g., the Red Cross) do not. *Non-financial time* incentives, however, are more commonly used to incentivize blood donation across Europe: 15 out of 28 countries offer some kind of time incentive. Countries that offer time off work to blood donors (independent of employer policies) include Eastern European countries such as Bulgaria and Romania, but also Southern European countries such as Italy and Portugal. In other European countries, a subset of employers allow donors to take time off work for making the donation (for instance in Belgium, Croatia, Greece and Sweden). Table S 2 provides detailed descriptions of financial and time incentives offered to blood donors across Europe.

Country-level rates of blood donation as a function of incentives and social norms Observational country-level mean blood donation rates as a function of the two country-level variables (a) incentives and (b) social

norms are displayed in Fig. S 6. Note that these are only raw observations, where no confounding factors are controlled for.

Empirical results: Statistical models

Full model results (including results from robustness checks) are displayed in Tab. S 3 (financial incentives) and Tab. S 4 (time incentives).

Country-level variation Models included random intercepts for country, which a log-likelihood test comparing the intercept-only model with and without random intercepts confirmed to be justified ($\chi^2(1) = 501.79, p < 0.001$).

Demographics Adding demographics to the model improved the model fit compared to the baseline (intercept-only) model further ($AIC_{demo} = 33097.4 < AIC_{intercept-only} = 34848.6$). As expected given that our dependent variable is blood donation *during the lifetime*, older respondents were more likely to have donated ($b = 0.166, p < 0.001$). The other socio-demographic control variables were also significantly related to blood donation in the following manner: male ($b = -0.471, p < 0.001$) and more educated respondents (full-time education until 16-19 years: $b = 0.395, p < 0.01$; full-time education until 20 years or more: $b = 0.697, p < 0.001$), and those living with a partner ($b = 0.129, p < 0.001$) were more likely to have donated blood during their lifetime. Predictive margins of donating blood as a function of norms and incentives, split by gender, are displayed in Fig. S 7.

Extrinsic motivation The effect of extrinsic motivation (in terms of getting something in return as a motivation to donate) was in the expected direction but non-significant ($b = 0.110, p = 0.078$ in models for financial incentives; $b = 0.104, p = 0.189$ in models for time incentives). Similarly, the interaction term between financial incentives and extrinsic motivation was in the predicted direction but did not reach significance (for financial incentive = 1: $b = 0.283, p = 0.09$; for time incentive = 1: $b = 0.108, p = 0.438$). In addition, the two other indicators with which we operationalized extrinsic motivation for *financial* incentives (i.e., lack of employment and difficulty paying bills) had no positive effect on blood donation, contrary to our expectations. Instead, lack of employment was associated with *lower* levels of blood donation and difficulty in paying bills was unrelated to donation behavior. In the models for *time* incentives we included *presence* of employment as an indicator of extrinsic motivation for time incentives, and indeed this had a positive effect on blood donation behavior ($b = 0.245, p < 0.001$).

Robustness checks We performed robustness checks for intrinsic and extrinsic motivation, because the Eurobarometer employed a selection criterion for questions regarding these two motivational variables: only respondents who had either donated in the past or were willing to donate in the future subsequently received the question concerning motivational factors. Since this selection criterion introduces significant bias which overestimates the effects of intrinsic and extrinsic motivation, we additionally ran all analyses only on those respondents who have donated blood or who indicated they are willing to donate in the future (i.e., excluding those indicating they are not willing to donate in the future). Results from these analyses are displayed in the rightmost columns in Table S 3 (financial incentives) and S 4 (time incentives).

Despite a sharp reduction in sample size and forgoing representativeness of the sample by excluding only a specific type of respondent, all effects found in the full model remained the same (except the demographic variable *cohabitation*, which turned non-significant). However, as expected due to the selection criterion, the effect of intrinsic motivation became considerably smaller ($b = 0.091, p < 0.05$).

Exploratory analyses: Time incentives for employed and unemployed individuals

We split the sample according to employment status and ran separate models for time incentives on these two subsets of respondents. While employed individuals should be motivated to receive a time incentive, those who are not employed cannot profit from time off work. However, the same results emerged for both of these models as for the main model for time incentives. This finding is likely due to temporal discrepancy of the measures (i.e., discrepancy between blood donation *during the lifetime* and *current* employment status; see limitations below).

Exploratory analyses: Region-level models

Lastly, exploratory analyses which investigated the association between within-country regional social norms and blood donation behavior revealed qualitatively similar patterns of results to our main confirmatory analyses (see supp. Fig. S 8A for financial incentives and S 8B for time incentives). However, uncertainty around the predictions increases, indicating that differences across social norms at the within-country region level (where regions may reflect cultural or mere geographical clustering) may be more subtle and thus cannot explain behavior as well as the country-level social norm measures.

Supplementary discussion

Obtaining (representative) samples from a sufficiently large number of countries is a major challenge for comparative studies such as ours. In the case of blood donation, the 2014 wave of the Eurobarometer is to our knowledge the only dataset which contains data on blood donation behavior that is comparable across multiple countries (as well as including attitudes towards different types of rewards and motivational factors). While this dataset thus has the advantage of offering rich population data for dozens of countries, our inability to precisely design the survey items has led to potential shortcomings regarding the operationalization of our theoretical constructs.

One issue concerns the Eurobarometer employing a selection criterion for questions regarding intrinsic and extrinsic motivation: only respondents who had either donated in the past or were willing to donate in the future subsequently received the question concerning motivational factors. As our chief interest concerned the relationship between incentives, social norms and blood donation (for which data was available for all respondents), this issue does not compromise our main results, which were computed based on the full sample (including donors and non-donors). However, the selection criteria does result in bias when examining the effects of intrinsic and extrinsic motivation. For this reason we performed robustness checks by running the main analyses on the subset of participants who received the question related to motivational factors; However, these robustness checks suffer from their own shortcomings, for instance decreased power due to starkly reduced sample sizes.

Moreover, it is problematic that the question regarding motivational factors included three intrinsic motivational factors as response options and only one extrinsic one, and that there may have been an order effect as the extrinsic motivational choice was always presented last. A measure of these motivational factors that targets more specifically individuals' pleasure in helping others and their utility of (non-)monetary rewards may be better able to quantify the effects of intrinsic and extrinsic motivation. Additionally, the Eurobarometer permitted only a very simplistic measure of social norms, namely the aggregation of acceptability ratings at the societal level. Particularly when there is lack of consensus in a society regarding the acceptability of an incentive, uncertainty about negative consequences to his reputation may lead the respondent to self-report disapproval of financial incentives, especially because this dishonesty is not costly to him (54, 47).

Additionally, some of our predictors may be less reliable due to suffering from a temporal discrepancy. That is, whereas our dependent variable is blood donation *during the lifetime*, many moderators are *current* indicators at the time of survey completion, for instance occupation and financial situation. Thus the current situation of the respondent may not correspond to the situation during which he donated. As an example of why this may be problematic, take Germany: One major group of donors that makes use of monetary incentives in Germany are university students (55), whose financial situation is relatively weak but who have good prospects of a high-paying position in the future. Lastly, we were not only constrained by the availability of *indicators* in the survey, but also by the number of *countries*: The inclusion of only 28 European countries limited the number of country-level indicators we could test in our multi-level models.

Supplementary methods

Variable operationalization

S 5 provides an overview of all variables included in the statistical models, including a detailed description, wording of survey items, and the response range.

Exclusion of respondents

For whole blood donation, 2.8% of the participants did not answer the dependent variable question on donation behavior and were excluded from the sample. Moreover, we excluded respondents younger than 18 years old, as these individuals are not eligible to donate in most European countries; this led to an additional 2.0% of respondents being excluded, resulting in a sample size of $n_1 = 26,532$.

In addition, the model for financial incentives excludes respondents from Germany, as Germany was the only country where *some* blood operators offer financial incentives. In all other countries financial incentives were offered either by all blood operators or by none. All results are quantitatively equivalent when Germany is included and regrouped either in the group of countries offering financial incentives or in the group not offering financial incentives. Note that for analyses pertaining to *time* incentives Germany is included (as multiple countries fall into each of the time incentive categories), which results in differing sample sizes for these two analyses (i.e., for *financial* incentives $n_{respondents} = 25068$, $n_{region} = 63$ and $n_{country} = 27$, whereas for *time* incentives $n_{respondents} = 26532$, $n_{region} = 68$ and $n_{country} = 28$).

Analytic strategy for exploratory region-level models

We followed the same analytic strategy for exploratory region-level analyses as we did for country-level analyses (i.e., including all individual-level variables such as intrinsic motivation, as well as *incentive* as a country-level variable). The key difference in the region-level analyses was that social norms (both for financial and time incentives) were aggregated at the NUTS level 1 region-level instead of at the country-level. NUTS (Nomenclature of Territorial Units for Statistics) is a commonly used geocode standard for referencing subdivisions of countries. Hereby NUTS *level 1* subdivisions are the next largest regional divisions after *country*. In order to ensure that samples drawn from these regions are sufficiently representative, we excluded NUTS level 1 regions from which less than 100 respondents were drawn. This resulted in the exclusion of eight regions from five different countries; analyses were run on $n = 26005$ participants from 86 regions within Europe. We included *regions* as additional random intercepts in all models. That is, the models assumed that participants are nested in regions, which again are nested in countries.

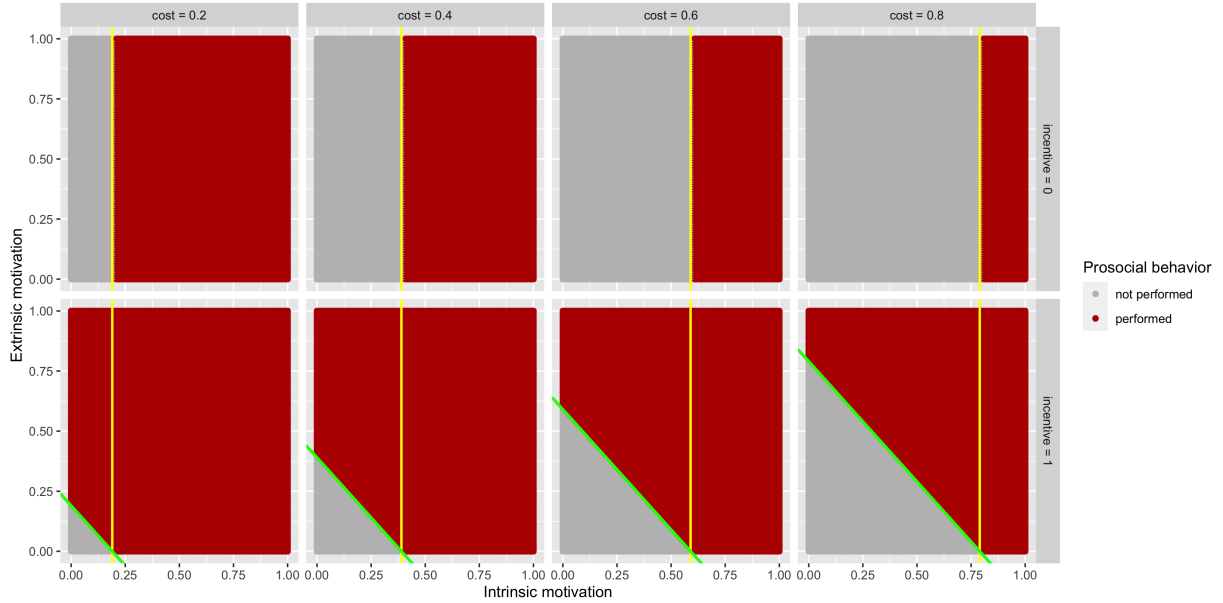


Figure S 1: Visualization of the effects of varying cost of the behavior (as depicted in the four columns) and the presence of incentives (as depicted in the two rows) for individuals varying in intrinsic and extrinsic motivation (where $v_a \sim \text{unif}\{0, 1\}$ and $v_v \sim \text{unif}\{0, 1\}$). The first row illustrates that in the absence of extrinsic rewards, intrinsic motivation and cost govern whether a behavior is performed (yellow line). For example, if cost is 0.4 and incentive = 0, then those individuals whose intrinsic motivation is greater than 0.4 perform the prosocial behavior, irrespective of their extrinsic motivation. In the second row, where incentives are provided, the line dividing those taking the action versus not (green line) shifts counterclockwise: More people engage in the prosocial behavior, because sufficiently extrinsically motivated individuals also engage in the behavior. For example, if cost is 0.6 and incentive = 1, those individuals whose intrinsic motivation is greater than 0.6 perform the prosocial behavior, as in the first row, and *additionally* those individuals whose sum of extrinsic and intrinsic motivation is greater than 0.6 perform the prosocial behavior.

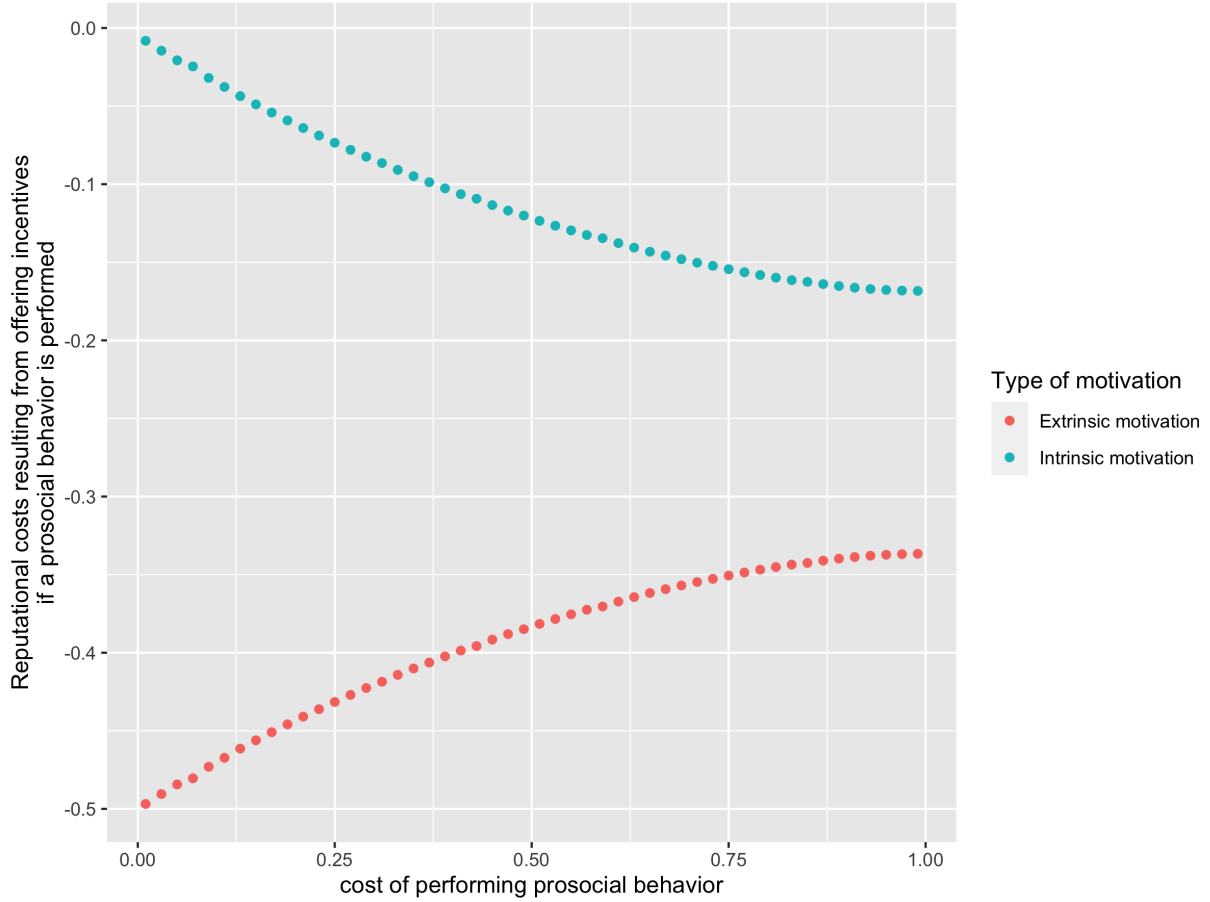


Figure S 2: Line plot displaying the additional reputational *cost* associated with performing a prosocial behavior in the presence of incentives, given BT’s assumption that higher levels of expected extrinsic motivation represent a cost and higher levels of intrinsic motivation represent a benefit. That is, the blue line represents the benefit in terms of higher expected *intrinsic* motivation when performing the behavior while receiving an incentive (i.e., $E(v_a|R = 1, B = 1) - E(v_a|R = 1, B = 0)$) **minus** the benefit in terms of higher expected intrinsic motivation when performing the behavior while *not* receiving an incentive (i.e., $E(v_a|R = 0, B = 1) - E(v_a|R = 0, B = 0)$). In other words, the blue line captures the *cost* of being perceived less intrinsically motivated when performing a prosocial behavior for which incentives are offered opposed to performing the behavior in the absence of incentives. The red line represents the cost in terms of higher expected extrinsic motivation when performing the behavior while receiving an incentive (i.e., $E(v_v|R = 1, B = 0) - E(v_v|R = 1, B = 1)$); note that it is not necessary to subtract the effect of expected extrinsic motivation when performing the behavior while *not* receiving an incentive, because without incentives there is no information about an agent’s extrinsic motivation, i.e. $E(v_v|R = 0, B = 0) - E(v_v|R = 0, B = 1) = 0$, see Eqn. 8).

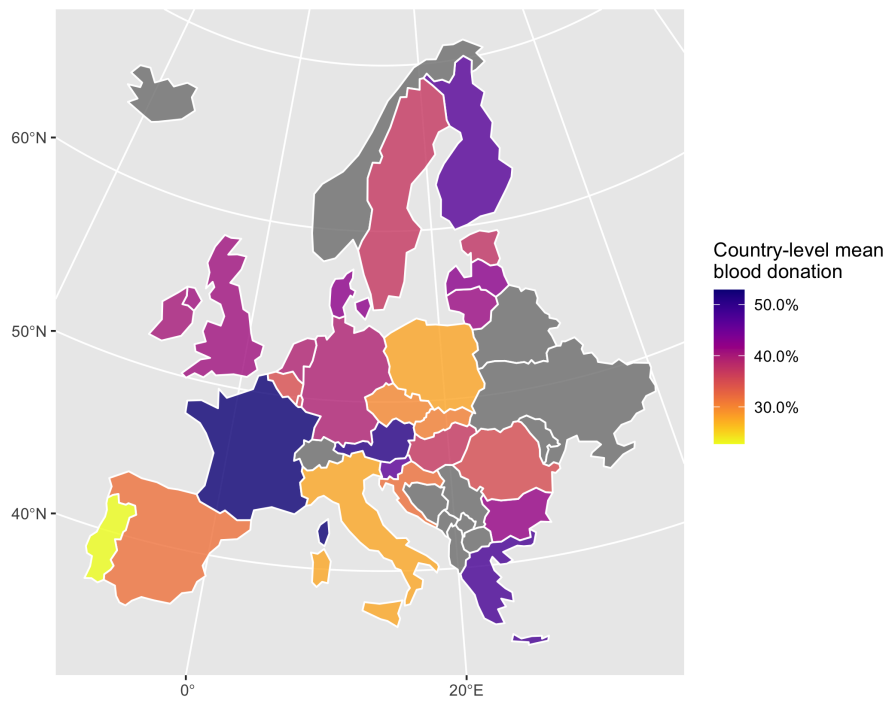


Figure S 3: Country-level mean levels of blood donation.

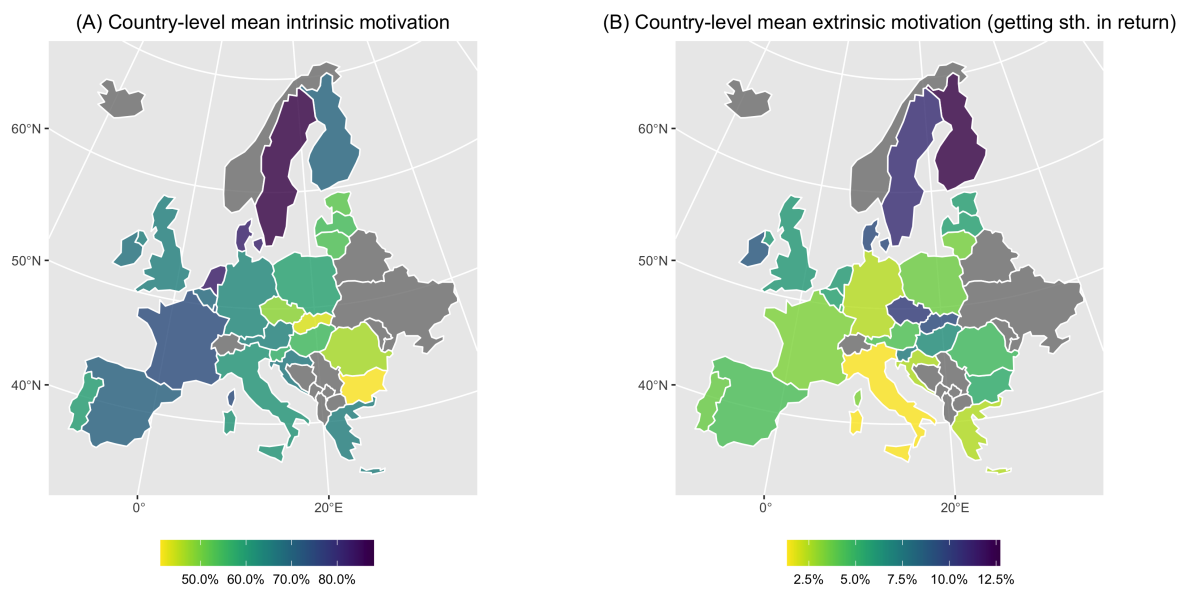


Figure S 4: Country-level mean levels of (A) intrinsic motivation and (B) extrinsic motivation (in terms of mentioning “getting something in return” as a motivation to donate).

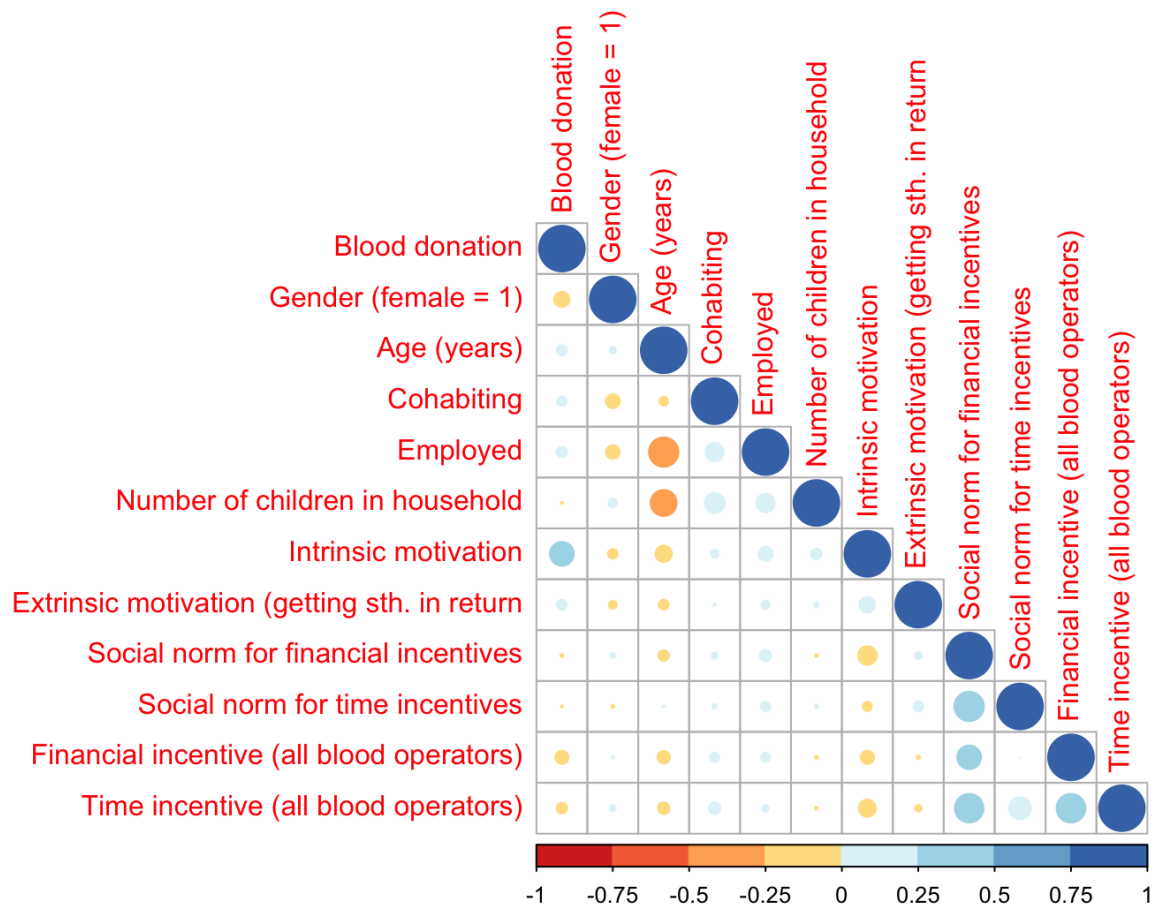


Figure S 5: Correlation matrix of predictors of interest (only continuous and dichotomous variables). Correlation coefficients involving a dichotomous variable reflect the point-biserial coefficient; otherwise coefficients are Pearson's correlation coefficients.

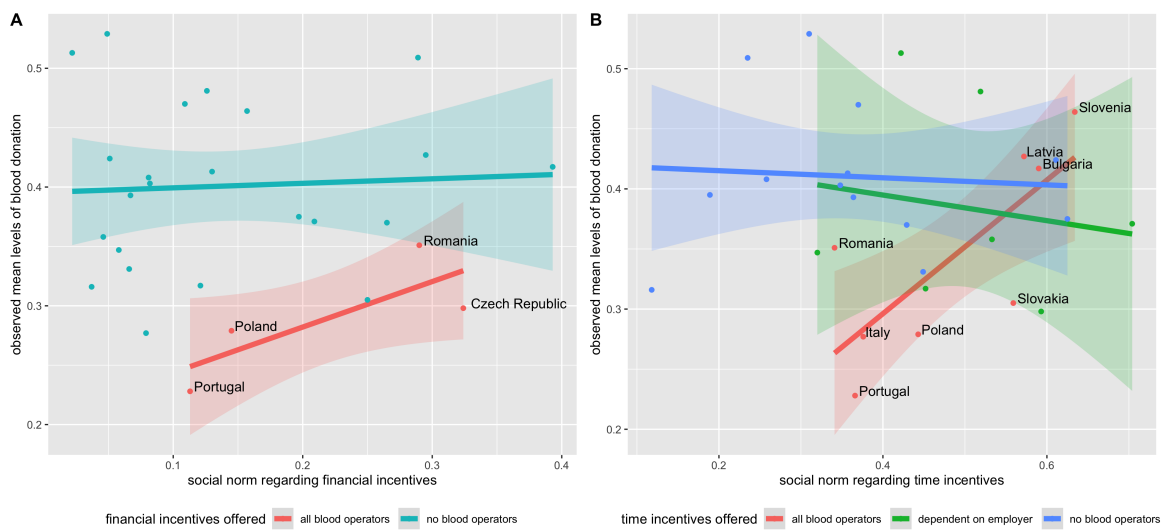


Figure S 6: Country-level mean donation rates as a function of (A) social norms regarding *financial* incentives (grouped by two levels of financial incentives) and (B) social norms regarding *time* incentives (grouped by three levels of time incentives).

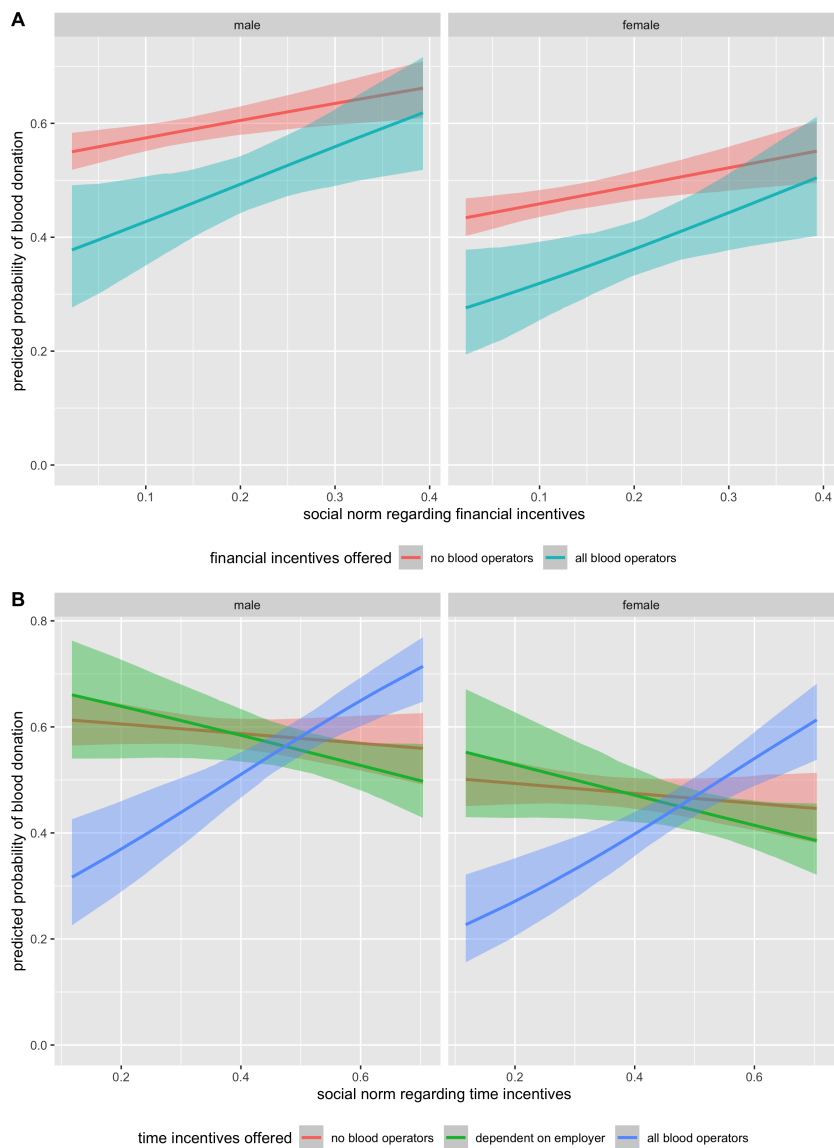


Figure S 7: Predictive margins of donating blood for a representative European man (left) and woman (right) (i.e., mean-aged, living with a partner, with secondary education, employed, living in a small town, with mean-number of children, rarely having difficulty paying bills, intrinsically but not extrinsically motivated) as a function of (A) social norms regarding *financial* incentives (grouped by two levels of financial incentives) and (B) social norms regarding *time* incentives (grouped by three levels of time incentives). Prediction bands indicate 80% prediction intervals.

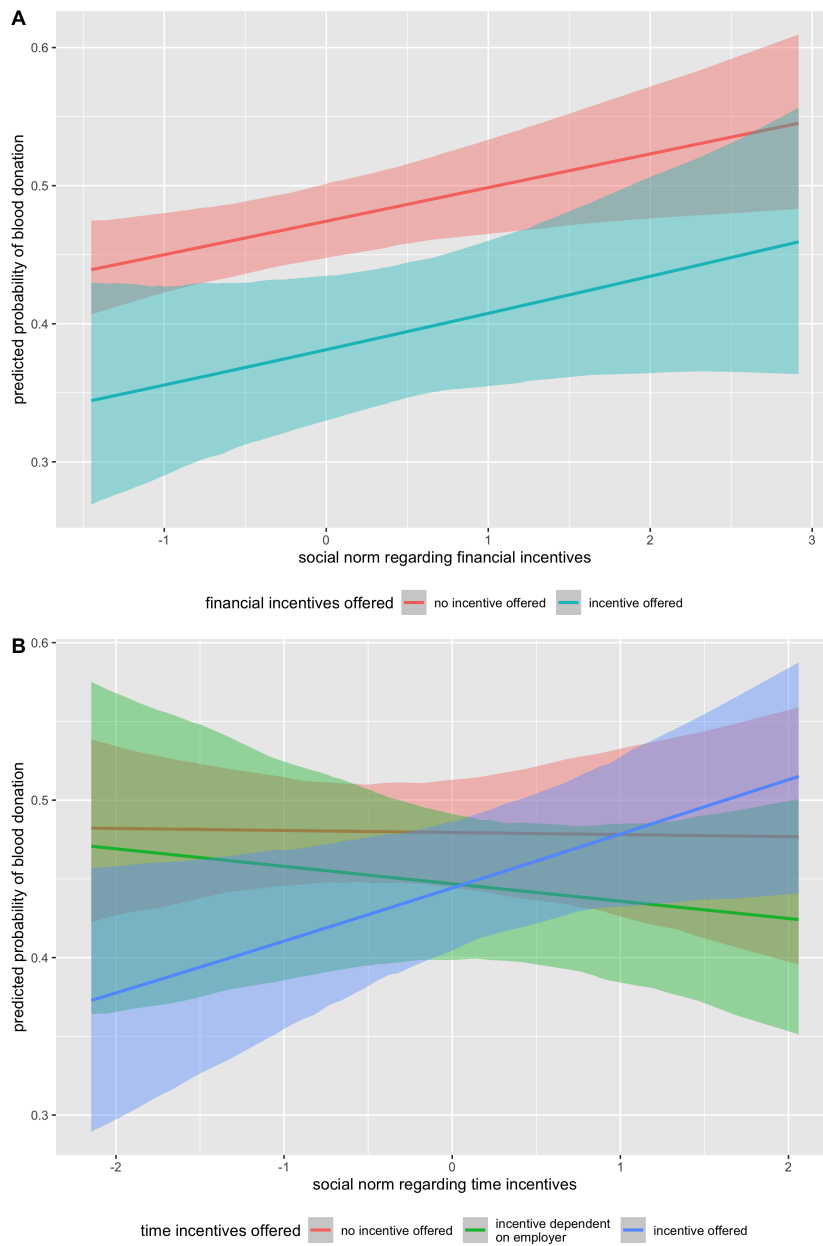


Figure S 8: Region-level exploratory analysis results: Predictive margins of donating blood for a representative European woman (i.e., mean-aged, living with a partner, with secondary education, employed, living in a small town, with mean-number of children, rarely having difficulty paying bills, intrinsically but not extrinsically motivated) as a function of **region-level** (A) social norms regarding *financial* incentives (grouped by two levels of financial incentives at the country-level) and (B) social norms regarding *time* incentives (grouped by three levels of time incentives at the country-level). Prediction bands indicate 80% prediction intervals.

Table 1: Descriptive statistics for the dependent and independent variables.

	Variable	Range	Mean	Std.	n
Individual-level	Blood donation	0; 1	0.38	-	26532
	Gender (female = 1)	0; 1	0.56	-	26532
	Age (years)	18-99	51.29	17.81	26532
	Education				
	None (ref.)	0; 1	0.01	-	25640
	≤ 15 years	0; 1	0.17	-	25640
	16-19 years	0; 1	0.45	-	25640
	≥ 20 years	0; 1	0.37	-	25640
	Cohabiting	0; 1	0.65	-	26476
	Employed	0; 1	0.49	-	26532
	Difficulty paying bills				
	Never (ref.)	0; 1	0.62	-	26192
	Sometimes	0; 1	0.27	-	26192
	Most of the time	0; 1	0.11	-	26192
	Intrinsic motivation	0; 1	0.63	-	26532
	Extrinsic motivation (getting sth. in return)	0; 1	0.06	-	26532
	Number of children in household	0-15	0.29	0.7	26527
	Type of community				
	Large town (ref.)	0; 1	0.27	-	26518
	Mid-sized town	0; 1	0.42	-	26518
Rural area	0; 1	0.31	-	26518	
Country-level	Financial incentives				
	No blood operators (ref.)	0; 1	0.82	-	28
	Some blood operators	0; 1	0.04	-	28
	All blood operators	0; 1	0.14	-	28
	Time incentives				
	No blood operator (ref.)	0; 1	0.46	-	28
	Dependent on employer	0; 1	0.25	-	28
	Independent of employer	0; 1	0.29	-	28
	Social norm (financial incentives)	0.02-0.39	0.15	0.1	28
	Social norm (time incentives)	0.12-0.70	0.43	0.15	28

Table 2: Description of financial and time incentives offered to blood donors in 28 European countries. 0 denotes *offered by no blood operators*, 0.5 denotes *offered only by some blood operators or only to some donors* and 1 denotes *offered by all blood operators*.

country	Financial incentive	Description	Time incentive	Description
Austria	0		0	
Belgium	0		0.5	depends on employers; up to 1 day off
Bulgaria	0		1	not depending on employer; 2 days
Croatia	0		0.5	depends on employers; time of donation
Cyprus	0		0.5	depends on employers; time of donation
Czech Republic	1	tax relief 17–18 Euro	0.5	depends on employers; duration of donation + travel
Denmark	0		0	
Estonia	0		0	
Finland	0		0	
France	0		0	
Germany	0.5	25 Euro cash (only some blood operators, not Red Cross)	0	
Greece	0		0.5	depends on employers, especially civil servants; 2 days
Hungary	0		0	
Ireland	0		0	
Italy	0		1	not dependent on employer
Latvia	0		1	not dependent on employer; 1 day
Lithuania	0		0	
Luxembourg	0		0.5	depends on employers; especially for public sector; duration of donation: ~1 hour
Malta	0		0	
Netherlands	0		0	
Poland	1	tax relief; 29,40 Euro/L	1	not dependent on employer; 1 day
Portugal	1	exemption from paying fees for access to the national healthcare (tax cut)	1	duration of donation: ~1 hour
Romania	1	food vouchers; 13 Euro	1	not dependent on employer; 1 day
Slovakia	0		1	not dependent on employer; up to 1 day
Slovenia	0		1	1 day off
Spain	0		0	
Sweden	0		0.5	depends on employers
United Kingdom	0		0	

Table 3: Model results for financial incentives (full sample and robustness check excluding respondents who indicated they are not willing to donate). Standard errors in brackets; *** denotes $p < 0.001$; ** denotes $p < 0.01$; * denotes $p < 0.05$; \diamond indicates variable was normalized.

	Variable	(1) Full sample	(2) Rob. check
Individual-level	(Intercept)	-1.302 (0.17) ***	0.029 (0.182)
	Age \diamond	0.312 (0.019) ***	0.394 (0.02) ***
	Gender (Female = 1)	-0.465 (0.028) ***	-0.457 (0.03) ***
	Cohabiting	0.079 (0.031) *	0.038 (0.033)
	Education		
	No full-time education (ref.)	-	-
	≤ 15 years	0.06 (0.154)	0.094 (0.165)
	16-19 years	0.289 (0.152)	0.234 (0.163)
	≥ 20 years	0.51 (0.152) ***	0.453 (0.163) **
	Not employed	-0.238 (0.033) ***	-0.209 (0.034) ***
	Difficulty paying bills		
	Never (ref.)	-	-
	Sometimes	0.027 (0.035)	0.041 (0.037)
	Most of the time	0.041 (0.051)	0.094 (0.055)
	Number of children in household \diamond	0.019 (0.015)	0.02 (0.015)
	Type of community		
	Large town (ref.)	-	-
	Mid-sized town	-0.046 (0.035)	-0.051 (0.037)
	Rural area	-0.091 (0.038) *	-0.099 (0.041) *
	Intrinsic motivation	1.346 (0.033) ***	0.091 (0.038) *
Extrinsic motivation (getting sth. in return)	0.11 (0.062)	-0.067 (0.062)	
Country-level	Financial incentive		
	No blood operator (0; ref.)	-	-
	All blood operators (1)	-0.512 (0.198) **	-0.485 (0.204) *
	Social norm for fin. incentives \diamond	0.125 (0.063) *	0.174 (0.065) **
	Social norm for fin. incentives $\diamond \times$ fin. incentive		
	Social norm \times incentive = 0 (ref.)	-	-
Social norm \times incentive = 1	0.137 (0.178)	0.099 (0.183)	
Cross-level interaction	Extrinsic motivation \times fin. incentive		
	Extrinsic motivation \times incentive = 0 (ref.)	-	-
	Extrinsic motivation \times incentive = 1	0.283 (0.168)	0.183 (0.167)
	N	23910	18961

Table 4: Model results for time incentives (full sample and robustness check excluding respondents who indicated they are not willing to donate). Standard errors in brackets; *** denotes $p < 0.001$; ** denotes $p < 0.01$; * denotes $p < 0.05$; \diamond indicates variable was normalized.

	Variable	(1) Full sample	(2) Rob. check
Individual-level	(Intercept)	-1.564 (0.177) ***	-0.168 (0.192)
	Age \diamond	0.316 (0.018) ***	0.394 (0.019) ***
	Gender (Female = 1)	-0.455 (0.027) ***	-0.445 (0.029) ***
	Cohabiting	0.076 (0.03) *	0.038 (0.032)
	Education		
	No full-time education (ref.)	-	-
	≤ 15 years	0.053 (0.151)	0.090 (0.163)
	16-19 years	0.258 (0.149)	0.200 (0.161)
	≥ 20 years	0.483 (0.149) **	0.419 (0.161) **
	Employed	0.244 (0.031) ***	0.213 (0.033) ***
	Number of children in household \diamond	0.021 (0.014)	0.022 (0.015)
	Type of community		
	Large town (ref.)	-	-
	Mid-sized town	-0.039 (0.034)	-0.044 (0.036)
	Rural area	-0.068 (0.037)	-0.075 (0.039)
Intrinsic motivation	1.369 (0.032) ***	0.093 (0.037) *	
Extrinsic motivation (getting sth. in return)	0.103 (0.079)	-0.033 (0.078)	
Country-level	Time incentive		
	No blood operator (0; ref.)	-	-
	Dependent on employer (0.5)	-0.033 (0.165)	-0.101 (0.179)
	Independent of employer (1)	-0.236 (0.151)	-0.201 (0.163)
	Social norm for time incentives \diamond	-0.055 (0.086)	-0.054 (0.093)
	Social norm for time incentives $\diamond \times$ incentive		
	Social norm \times incentive = 0 (ref.)	-	-
Social norm \times incentive = 0.5	-0.116 (0.169)	-0.113 (0.183)	
Social norm \times incentive = 1	0.483 (0.168) **	0.479 (0.182) **	
Cross-level interaction	Extrinsic motivation \times time incentive		
	Extrinsic motivation \times incentive = 0 (ref.)	-	-
	Extrinsic motivation \times incentive = 0.5	0.068 (0.144)	0.092 (0.142)
	Extrinsic motivation \times incentive = 1	0.107 (0.139)	-0.111 (0.137)
	N	25586	20218

Table 5: Overview of dependent and independent variables. ▷ denotes source: Eurobarometer; ★ denotes source: Expert survey.

Variable	Description	Original item	Response range
Blood donation	Individual-level response: Blood donation during lifetime ▷	During the lifetime of a person it is possible to donate different body substances (blood or cells) to help other people. Could you please indicate which ones you have [...] donate[d] yourself? Blood	0; 1 (dummy coded) 0 (not donated blood in the past) 1 (donated blood in the past)
Intrinsic motivation	Donation motivated by intrinsic factors (i.e., help people in need, alleviate shortages, medical research) ▷	You said that you have or would be prepared to donate certain body substances during your lifetime (blood or cells) or after death (tissues). For which of the following reasons have you or would you donate any of the body substances mentioned earlier? To help other people in need To alleviate shortages of these substances To support medical research	0; 1 (dummy coded) 0 (did not mention intrinsic motivator) 1 (mentioned intrinsic motivator)
Extrinsic motivation	Donation motivated by extrinsic factors (i.e., to receive something in return) ▷	You said that you have or would be prepared to donate certain body substances during your lifetime (blood or cells) or after death (tissues). For which of the following reasons have you or would you donate any of the body substances mentioned earlier? To receive something in return for you or your relatives	0; 1 (dummy coded) 0 (did not mention extrinsic motivator) 1 (mentioned extrinsic motivator)
	Subjective difficulty to pay bills ▷	During the last twelve months, would you say you had difficulties to pay your bills at the end of the month...?	0; 1; 2 (dummy coded) 0 (Never/Almost never) 1 (From time to time) 2 (Most of the time)
	Lack of current occupation ▷	What is your current occupation?	0; 1 (dummy coded) 0 (other) 1 (Homemaker, student, unemployed, retired or unable to work through illness)
Personal cost	Time constraints due to child care ▷	Could you tell me how many children less than 10 years old live in your household?	numeric
	Travel costs due to rural living ▷	Would you say you live in a... Rural area or village Small or middle sized town Large town	0; 1; 2 (dummy coded) 0 (lives in large town) 1 (lives in small or middle sized town) 2 (lives in rural area or village)

Social norms	Aggregated response at country-level: Mean social norm regarding financial incentives for blood donation ▷	For donating blood or plasma during someone's lifetime, do you consider it acceptable to receive cash amounts additional to the reimbursement of the costs related to the donation?	[0, 1] (aggregated from individual-level $i \in \{0; 1\}$) 0 (0% of population found acceptable) 1 (100% of population found acceptable)
	Aggregated response at country-level: Mean social norm regarding time incentives for blood donation ▷	For donating blood or plasma during someone's lifetime, do you consider it acceptable to receive time off work (for the time needed for the donation and/or for recovery)	[0, 1] (aggregated from individual-level $i \in \{0; 1\}$) 0 (0% of population found acceptable) 1 (100% of population found acceptable)
Controls	Age of respondent ▷	How old are you?	numeric
	Gender ▷	NA	0; 1 (dummy coded) 0 (Male) 1 (Female)
	Education (Age when finished full-time education) ▷	How old were you when you stopped full-time education?	0; 1; 2; 3 (dummy coded) 0 (No full-time education) 1 (15 years or younger) 2 (16-19 years) 3 (20 years and older)
	Partner status (live with partner or not) ▷	Could you give me the letter which corresponds best to your own current situation? (recoded from 14 levels, e.g., married/remarried and living with/without children)	0; 1 (dummy coded) 0 (Living without partner) 1 (Living with partner)
Incentives	Country-level variable for whether financial incentives are offered for donating blood. ★	See expert survey	0; 0.5; 1 (dummy coded) 0 (no blood establishments offer financial incentives for donating blood) 0.5 (some blood establishments offer financial incentives for donating blood) 1 (all blood establishments offer financial incentives for donating blood)
	Country-level variable for whether time incentives are offered for donating blood. ★	See expert survey	0; 0.5; 1 (dummy coded) 0 (no time incentives are offered for donating blood) 0.5 (time incentives are offered for donating blood dependent on employer) 1 (time incentives are offered for donating blood independent of employer)

Supplementary references

54. A. Diekmann and P. Preisendörfer, “Green and greenback: The behavioral effects of environmental attitudes in low-cost and high-cost situations,” *Rationality and Society*, vol. 15, no. 4, pp. 441–472, 2003. Publisher: Sage Publications
55. C. Weidmann and H. Klüter, “Blood collection and donor motivation in Germany,” *ISBT Science Series*, vol. 8, no. 1, pp. 238–241, 2013. Publisher: Wiley Online Library.