WORKING PAPERS

The Roman Metro Problem

Christian Roessler
Sandro Shelegia

March 2012

Working Paper No: 1202

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

The Roman Metro Problem

Christian Roessler and Sandro Shelegia
University of Vienna

March 1, 2012

Abstract

In Rome, if you start digging, chances are you will find things. We consider a famous complaint that justifies the underdeveloped Roman metro system: "if we tried to build a new metro line, it would probably be stopped by archeological finds that are too valuable to destroy, so we would have wasted the money." Although this statement appears to be self-contradictory, we show that it can be rationalized in a voting model with diverse constituents. Even when there is a majority preference for a metro line, and discovery of an antiquity has the character of a positive option, a majority may oppose construction. We give sufficient conditions for this inefficiency to occur. One might think it arises from the inability to commit to finishing the metro (no matter what is discovered in the process). We show, however, that the inefficient choice is made in voting over immediate actions precisely when there is no Condorcet winner in voting over contingent plans with commitment. Hence, surprisingly, commitment cannot really solve the problem. Our results extend to other common dynamic voting scenarios, such as the academic job market, which share the essential features of the Roman metro game.

JEL *classification*: D70, H41, C70.

1 Introduction

The Roman underground has two lines, 49 stations, that serve a metropolitan area of 3.4 million residents and 9 million annual visitors. Berlin is similar in size, but has 173 subway stations. Madrid, about one-and-a-half times as big as Rome, has 300. Even in Oslo, where less than a million people live, the metro has 105 stops.[1] What is lacking underground cannot be compensated on the surface - the eternal city was not built with suspended monorails and large buses in mind. So why don't Romans invest more in their metro system? Every Roman will eagerly explain why: metro projects are extremely expensive, and in Rome, more than anywhere else in the world, one is likely to run into ancient ruins of such value that the metro project would be suspended to preserve the ruins.[2] Hence, Romans stopped trying to build metro lines altogether.

[1] London and Paris, with 8.6 and 10.5 million people (according to the UN definition), have respectively 270 and 301 stations. There is no major metropolitan area in Europe that has more inhabitants per metro stop than Rome - except Athens, which has 3.3 million inhabitants and 33 metro stops. It's no coincidence ... as the paper proceeds to explain.

[2] The issue is actually not the metro tunnel itself, which could be built below the archeological layer; it's the various vents and exits that have to be dug at regular distances.

There is only one small problem with this explanation: on the face of it, this argument makes little sense. Consider the following analogy. Your house needs water so you want to dig a well in a place where you suspect there may also be gold. It is possible that you will strike upon gold while excavating, get rich and never finish the well. Given this "risk," you decide not to dig in the first place. It seems like a ridiculous conclusion to draw because the "risk" has no downside, it's an option that might materialize, and options cannot have negative value (you can, after all, drill right past the gold if you want to). With the possibility of finding gold, digging the well is more, not less, attractive - either your house gets water supply or you get rich and move out.

From this standpoint, Romans should not worry that they might discover ancient ruins. What happens if they do not build the metro line? They will never know there are ruins in the way.[3] If they find ruins, they can still destroy them. If they do not wish to destroy them, then apparently the ruins are more valuable than the metro - and the construction project has paid off more than initially expected. Romans should hope they find ruins! It's a valuable option.

Or is it that simple? The above argument implicitly presumes that there is a single decision maker. Only a time-inconsistent person would refuse to start a project for fear of not finishing it because she might pursue an even better opportunity that appears in the process of implementing the initial project. We show that such time inconsistency can easily arise as a result of a political process with diverse constituents.[4]

We study the Roman metro problem as a majority voting game in two stages. If an initial majority decision is made to build a metro, an antiquity will be found with some probability, which triggers a second vote to determine whether construction should continue. Suppose (i) a majority values the metro more than the cost of construction; (ii) the total utility from metro construction is also positive; (iii) a majority values the antiquity that may be discovered more than the metro; (iv) the total utility from the antiquity is also higher than the total utility from the metro. Given (iv) the antiquity represents a positive option for a social planner, which in addition to (i) means that the social planner would start the metro project. We show that nevertheless a majority of voters might oppose beginning construction, which is inefficient.

Specifically, there normally (under plausible restrictions on the payoff distribution) exists a range of probabilities of finding an antiquity that cause a majority to oppose construction. Some voters will reject the metro outright because they find it too costly, but would favor preserving an antiquity if found (when construction has already begun and cost is sunk). Metro proponents must anticipate that some who are initially in their

[3] A similar sentiment was voiced by Enrico Testa, the chairman of Roma Metropolitane SpA, who said in 2007: "There are treasures that are underground that would stay buried forever, but as soon as we uncover them, our work gets blocked." (Appeared in the Wall Street Journal, January 27, sec. A, p. 1.)

[4] The immediate causes of Rome's inability to develop its metro system do not necessarily have the appearance of a voting game. Romans might mention misaligned interests between those in charge of metro construction and the city dwellers. Or the independence of the archeological authority, the Sopraintendenza ai Beni Culturali (Superintendency of Cultural Heritage), which can hold the city hostage to its devotion for the preservation of antiquities that ordinary Romans would sacrifice for the metro. We take the view, as is common in political economy analysis, that the existence and decisions of institutions ultimately have to reflect majority opinion. The underlying policy preferences factor in local aspects such as the inefficiency of public works, disruptions from construction in a densely populated city like Rome, etc.

camp will join these metro opponents in voting to preserve the antiquity, if it is discovered. Those who value the antiquity less than the cost of the metro line may then vote against their inclination to build the metro. Thus, shifting coalitions lie at the heart of our argument.

Many authors have highlighted that a Condorcet cycle among three or more policies may lead to strategic voting in a dynamic setting. In the Roman metro problem, all votes are between two actions, so that a Condorcet cycle cannot occur. This is because we rule out commitment to future choices. In a hypothetical world with commitment, there is a richer set of alternatives - (i) take no action, (ii) start building the metro and continue regardless of whether an antiquity is found or not, and (iii) start building the metro but stop if an antiquity is discovered. It turns out that a Condorcet cycle on this set is necessary and sufficient for a majority to vote against construction in our model without commitment. Therefore, commitment does not help eradicate economic inefficiency because it leads to a situation without a Condorcet winner. That the failure to implement a beneficial project can ultimately be linked to a Condorcet cycle, even though the game is quite different in character, suggests that the relationship between Condorcet cycles and economically inefficient majority choices may be a fundamental one.

We are not aware of research that has examined the precise game we model, or classes that contain it, but there are various literatures on voting games with related features. We contrast them briefly. In models with dynamic Condorcet cycles (e.g. Penn (2009); see also Roberts (2007), and Bernheim and Slavov (2009)), a status quo is pitted against a new alternative in each period, and payoffs accrue in every period. Voters not only consider their immediate benefit from the current policy, but also future votes and their outcomes. Since the presently preferred policy may be a weak competitor against an undesirable alternative that may come up next, voters have to be "farsighted" and potentially vote against an option with a higher immediate payoff. In our model, agents do not encounter the same trade-off, since there are no payoffs beyond the ones arising from the final outcome. Instead, a change in voter alliances, triggered by a random event (finding an antiquity) can dissolve a majority preference (for building the metro) that exists today. Metro supporters anticipate a possible future vote which will terminate the project and, because the costs are sunk, may not be willing to take the risk.

Barbera et al. (2001) and Jack and Lagunoff (2006) study situations where current members of a group vote to include new members who will get voting rights if admitted. Agents may vote strategically for less preferable newcomers, anticipating that their voting behavior in the future will lead to favorable outcomes. These models are related to our model in that the voting blocks change over time. In our model, a group that gets no net benefit from either the metro or the antiquity would like to be excluded from the project, but because this is not possible, its potentially pivotal role in preserving the antiquity, induces a group that wants the metro to reject it.

There is a literature on voting rules that are implementable by backward induction, also called sophisticated voting (see McKelvey and Niemi (1978), Srivastava and Trick (1996), and Dutta and Sen (1993), among others). The objective is to narrow a set of alternatives down to a single choice through pairwise majority voting, as in an elimination tournament. Hence, the voting process starts at the leaves and ends at the root with a winner. The focus is on the implementability of voting rules on suitably chosen trees. We, on the other hand, are concerned with the efficiency of voting on a specific tree that reflects a particular decision problem, where voting proceeds from the root to the leaves. Our tree does not represent a voting agenda where the aim is to choose a single policy, but rather it is a conventional game tree where at some nodes a majority needs to make a decision.

We begin with an example of how the economically efficient choice, to start construction, can fail in the Roman metro problem. Then we present graphic intuition for how the example generalizes and demonstrate that majority decisions to build or not build the metro (without commitment to what happens in case an antiquity is found) are determined by the existence of a Condorcet cycle under commitment. Then, we derive restrictions on the payoff distribution, which guarantee respectively a yes or no vote on construction (or equivalently the absence or presence of a Condorcet cycle under commitment). Finally, we illustrate that the Roman metro game is also applicable to a variety of other dynamic voting environments.

2 An Example

The city has the option to build a metro, but there is a probability q that, while doing so, an antiquity will be found that lies in the path. If so, the antiquity could either be destroyed, or the construction project could be abandoned. The decision maker (mayor) is acting in line with majority preferences.[5]

Part of the consideration in building the metro is that it will have to be financed (through taxes or cuts elsewhere). Therefore, negative valuations are entirely possible. Suppose that the population consists of three types (A, B, C) of equal mass, so any two types together form a majority. A wants to build the metro, but only if it will not be abandoned if an antiquity is found. B wants to build the metro and abandon it if an antiquity is found. C does not want to build the metro and, if it is built and an antiquity is found, wants to abandon it. Thus, we have the following game tree, with some possible payoffs that reflect the preferences of the three constituencies:

Note that, if $q = 0$ (i.e. there is no possibility of finding an antiquity), building the metro has majority support (from A and B) and yields a surplus of 1/3 per voter. If $q > 0$, and an antiquity is found, a majority (B and C) will vote in favor of abandoning the metro. Abandoning the metro not only has majority support, but also is an option that is strictly valuable in terms of social surplus (per voter utility from abandoning is 2/3 while it is 1/3 from metro).

Anticipating future voting outcomes, A initially votes for construction if

$$-2q + 1 - q \geq 0 \iff q \leq \frac{1}{3}.$$

[5]See footnote 4 for a discussion of this assumption.

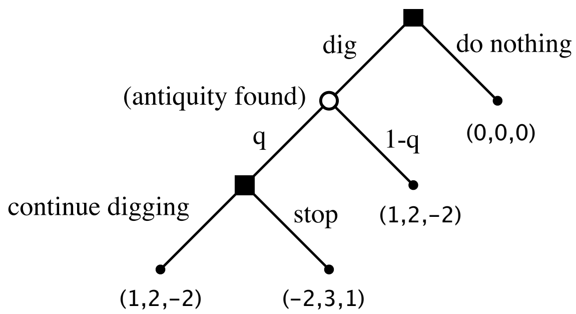


Figure 1: The Roman Metro Game

B votes for construction regardless of q, because it benefits whether or not an antiquity is found. C votes for construction if

$$q - 2(1-q) \geq 0 \iff q \geq \frac{2}{3}.$$

Overall, there is a majority for construction at the outset if $q \leq \frac{1}{3}$ (in which case, the project is supported by A and B) or $q \geq \frac{2}{3}$ (then, the project is supported by B and C). But in case $\frac{1}{3} < q < \frac{2}{3}$, a majority (consisting of A and C) will oppose digging.

At the same time, for any q, the expected utility per voter from the project is $(2/3)q + (1/3)(1-q)$. The latter is always positive, so a social planner would start the project regardless of q, and stop if an antiquity were found. Hence, adding the option to abandon may lead to an economically inefficient decision if q is intermediate. Rome's problem is that the probability of finding an antiquity is relatively high (while most other cities can ignore the possibility), yet not high enough for those who would like to discover an antiquity to support digging only as a way to search for antiquities. As we will see, there is a range of probabilities q that leads to the inefficient choice in this game under more general conditions.

3 The General Problem

We define the Roman metro problem as a voting game where the sequence of decisions is as in the example above, but the population consist of arbitrary types, and payoffs satisfy some restrictions that stack the deck squarely in favor of digging. We use M to denote the outcome that the metro is completed and T for the preservation of the antiquity (with the consequence that the metro line is never finished). The lower-case letters (m and t) represent the associated payoffs.[6]

The initial choice is either Y ("yes") to start digging or N ("no") to shelve the project. The decision tree, from the point of view of an individual citizen, looks as in Figure 2. At

[6] Payoff t from abandoning the metro for the antiquity could be the expectation of a lottery over various stages at which the antiquity may be found. Finding it earlier may mean that less cost is sunk into building the metro, and this would increase the effective value of the antiquity. Specifically, if discovery happens at a stage where only some share $\lambda < 1$ of the total cost c of the metro is sunk, the value of the antiquity increases for all types to $t + (1-\lambda)c$. Then, a type votes for digging if $(1-q)m + qt + q(1-\lambda)c) \geq 0$. This is a weaker inequality to satisfy than previously.

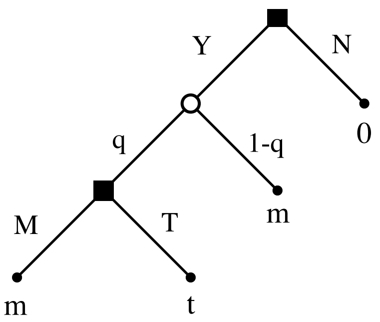


Figure 2: Decision Problem

the square nodes, decisions are made according to voting majorities. The circular node is a chance node. Action Y (dig) leads either to M or, with probability q, a choice between M and T. Inaction N gives a payoff of zero.

We impose that T beats M, and M beats 0, both in terms of majority preference and average payoff. Hence, for at least half the population,

$$t > m$$

and, for at least half the population,

$$m > 0.$$

(a majority prefers the metro over nothing, and another majority prefers the antiquity over the metro). Moreover

$$E(t) > E(m) > 0$$

(on average, across the population, the metro yields a positive payoff, and the antiquity yields a greater payoff than the metro). Therefore, from a social standpoint, the metro is desirable, and the antiquity is a valuable option if found. This assumption is not implied by the previous two about the majority support. It could have been that a majority supports the antiquity over the metro even though the average payoff of the metro is higher. By assuming that, in the pairwise comparisons, a majority sides with the average voter (the social planner), we bias the game against the possibility of an inefficient decision on the project.

It is crucial that the majority that favors m over 0 can have a different composition from the majority that prefers t to 0. If a majority of citizens individually preferred both m and t to 0, the initial vote would clearly be in favor of construction. Since we are interested in the possibility that a majority opposes building the metro, we rule this out: those who prefer both m and t to 0 constitute less than half the population. Let there be a distribution of types θ, who are described by a payoff tuple (m_θ, t_θ). Then, the above statements amount to restrictions on the type distribution, which we refer to (in combination with the game form) as the Roman metro problem.

To summarize our assumptions: a majority (and the social planner) prefers building the metro over doing nothing, a majority (and the social planner) prefers the antiquity over the metro, but those who prefer both the metro and the antiquity to doing nothing are in the minority.

4 Graphic Analysis and Condorcet Cycle Characterization

In order to find the expected payoff from digging for any agent, recall that the antiquity by assumption, if found, has majority support over the metro, hence the project stops. This means that the project is a lottery where the agent receives a payoff of m with probability $1-q$ and t with probability q. The payoff from doing nothing is assumed to be zero, so the individual supports digging if

$$(1-q)m + qt > 0,$$

and is against starting the project otherwise.

Now we turn to efficiency. A social planner who wants to maximize total payoff has a simple problem to solve. If an antiquity is found, he should choose the option with the highest total payoff. Since $E(t) > E(m)$ by assumption, he should stop building the metro in that case. This gives expected welfare

$$\begin{aligned} & (1-q)E(m) + q\max\{E(m), E(t)\} \\ = \quad & (1-q)E(m) + qE(t) \end{aligned}$$

from the project. The expression $\max\{E(m), E(t)\}$ reflects that the antiquity is an option that does not need to be exercised. The social planner's implied decision rule

$$(1-q)E(m) + qE(t) > 0$$

is identical to that of an individual agent, with the sole difference that the social planner uses average payoffs to make the decision.

Figure 3 depicts our opening example for $q = \frac{1}{2}$. The line $(1-q)m + qt = 0$ represents the types who are indifferent between digging and doing nothing. This line has a slope of $-(1-q)/q$ and passes through the origin. Agents whose payoff pairs (m, t) lie above the indifference line will vote for the project, all others will vote against it. The line $m = t$ separates agents who prefer the antiquity to the metro from those who prefer the metro to the antiquity. Each type is shown by a dot that reflects valuations for the metro and the antiquity. The dot labeled P depicts the average payoffs used by the planner, who can be thought of as an agent with payoffs $(E(m), E(t))$.

A two-thirds majority (groups B and C) prefers the antiquity to the metro, another two-thirds majority (A and B) favors the metro over doing nothing, but the project only has the support of a third of the population (B). At the same time, average payoffs from the antiquity and the metro are such that the social planner would implement the project. Hence, even though votes are aligned with the social planner's choice in the direct comparisons of metro vs. antiquity and metro vs. doing nothing, the majority is against a socially desirable project.

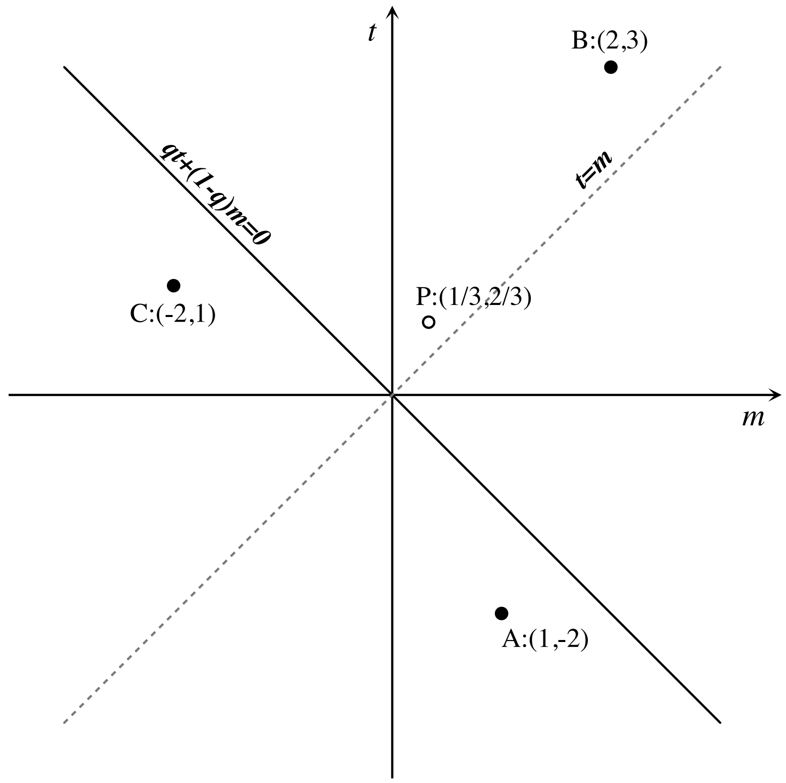


Figure 3: Majority Separated from Average Type

Such a situation is not unique. All that is required of the distribution of preferences is that a majority is below the indifference line, while the average type is above the line. In fact, given any distribution of types, if there exists a downward-sloping line through the origin that separates the average type from the majority of the population, then there exists q such that the majority takes the socially inefficient decision not to dig.

To see this, assume that we can find such a line, so that $t = -\beta m$ for some $\beta > 0$. If we set $q = 1/(1+\beta)$, the line $t = -\beta m$ is the same as the line $t = -((1-q)/q)\,m$, which is comprised of the types who are indifferent between digging and not digging. But since the social planner and the majority are on opposite sides of this line, they support different choices; thus, the majority decision cannot be efficient in terms of total surplus.

Visual inspection shows that getting a majority to reject the project does not require a very special type distribution. On the other hand, symmetric distributions, such as a uniform distribution on a circle, will not produce the inefficiency - in such cases, the average type and the majority always vote the same way.[7] As q goes to zero, we always get to a socially efficient decision, since the expected value from digging approaches $E(m) > 0$ from above, and the majority prefers the metro to doing nothing, hence will support the project if the chance of finding an antiquity is small enough. As Figure 4 illustrates, if in our example we decrease q from $\frac{1}{2}$ to $\frac{1}{4}$, group A swings from opposing to supporting the project, and the project is majority-approved.

[7] See Proposition 4, later in the text.

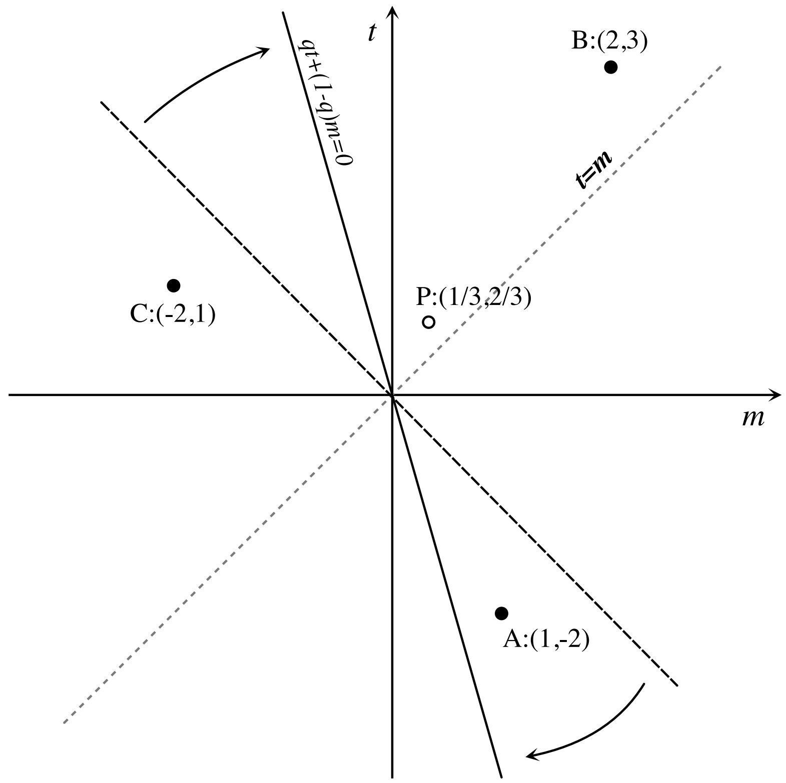


Figure 4: Majority Aligned with Average Type

4.1 Condorcet cycles

In the Roman metro problem, alternatives are defined such that one cannot commit beforehand to a course of action; decisions are only over the "next step." If the contingent plans (YM) "dig and build the metro, whether or not an antiquity is found," (YT) "dig, but preserve the antiquity if found" and (N) "do nothing" were put to the vote, then, in our initial example, A could condition support for the project on a commitment to continue even if an antiquity is found. It would then be joined by B in favoring digging over doing nothing. (The fully efficient choice, to dig but preserve the antiquity, is only possible if q is large enough so that C will join B in support of it.) It turns out that majority opposition to digging without commitment is equivalent to the existence of a Condorcet cycle over contingent plans.

Proposition 1. *In the Roman metro problem, a majority opposes digging if and only if there exists a Condorcet cycle over the set of contingent plans that would be feasible with commitment.*

Proof. If a Condorcet cycle exists, then it has to take the form that (YM) is majority-preferred to (N), (YT) is majority-preferred to (YM) and (N) is majority-preferred to (YT). The reason is that (YM) delivers m with certainty, and by assumption, for a majority $m > 0$, where 0 is the certain outcome of (N). Moreover, (YT) is a lottery between m and t with expected value $(1-q)\,m + qt$, which is majority-preferred to m because $t > m$ for a majority. The only way to get a Condorcet cycle is then for (N) to beat (YT) in majority voting. But this is precisely the choice voters make in the Roman metro problem when they forego digging, because $(1-q)\,m+qt\ \leq\ 0$ for a majority. Conversely, if a majority opposes digging in the Roman metro problem, then it favors 0 over $(1-q)\,m+qt$, i.e. (N) is majority-preferred to (YT). Because (YT) is majority-preferred to (YM), and (YM) is majority-preferred to (N) by the assumptions of the game, we have a Condorcet cycle. ■

In our example, we found that the majority vote against digging is supported by $q \in \left[\frac{1}{3}, \frac{2}{3}\right]$. With the probability of finding an antiquity in this interval, A and C oppose the metro project. When is there a Condorcet cycle between the contingent plans (YM), (YT) and (N)? When (N) beats (YT), i.e. $(1-q)\,m+qt\ \leq\ 0$ for a majority. This majority must consist of A and C, since B prefers both metro and antiquity to doing nothing. But A and C both satisfy the inequality exactly when $q \in \left[\frac{1}{3}, \frac{2}{3}\right]$. Thus, the requirement for the Condorcet cycle to exist is the condition for a majority to oppose digging.

It turns out then that it is not simply the inability to commit to contingent plans that causes inefficiency in the Roman metro problem. If commitment is possible, and there is a Condorcet winner among contingent plans, a majority supports construction of the metro even without commitment, i.e. when voting on immediate actions.[8] Whenever a majority would oppose construction, there is no Condorcet winner among the contingent plans, so commitment does not guarantee an efficient outcome. Rather, we must look at the distribution of preferences to understand the source of inefficiency.

5 Sufficient Conditions on the Type Distribution

We consider now properties of the type distribution (beyond the ones which are already imposed by the game form), which guarantee either that the metro project fails or succeeds in majority voting. By the equivalence we just established, these conditions are also sufficient for a Condorcet cycle on contingent plans (involving commitment) to exist or not exist. Let u^1_θ denote type θ's utility from the top-ranked alternative, u^2_θ the utility from the mid-ranked alternative, and u^3_θ the utility from the bottom-ranked alternative.

Definition: Type θ satisfies *decreasing differences* if $u^1_\theta - u^2_\theta \ \leq\ u^2_\theta - u^3_\theta$. Type θ satisfies *increasing differences* if $u^1_\theta - u^2_\theta \geq u^2_\theta - u^3_\theta$.

That is, with decreasing differences, the top-ranked alternative improves on the mid-ranked alternative by *less* than the mid-ranked alternative improves on the bottom-ranked alternative. In other words, moving from the worst to the best alternative, one gains smaller increments of utility. With increasing differences, the situation is opposite. Moving from the worst to the best alternative, one gains larger increments of utility.

[8]In the absence of a Condorcet cycle on contingent plans, there is a majority for digging (and preserving the antiquity), regardless of whether agents are voting over contingent plans or immediate actions. Construction will start and terminate in case an antiquity is found, in line with efficiency.

Though such properties might conceivably hold for the entire population, for the purpose of our results only a critical subset of types needs to have them. Let θ be a *pivotal type* if $u^2_\theta = 0$, or expressed differently, $\max\{m_\theta, t_\theta\} \geq 0 \geq \min\{m_\theta, t_\theta\}$. That is, θ neither prefers both M and T to zero, nor prefers zero to both M and T. This means that θ's preference between the lottery that yields $(1-q)\,m + qt$ and zero depends on the magnitude of q.[9] A pivotal type satisfies decreasing differences if and only if $\max\{m_\theta, t_\theta\} \leq -\min\{m_\theta, t_\theta\}$, and positive differences if and only if $\max\{m_\theta, t_\theta\} \geq -\min\{m_\theta, t_\theta\}$.[10] Starting at zero (inaction), with decreasing differences, I lose more from switching to my least-preferred option than I gain from switching to my most-preferred option. With increasing differences, it is opposite. Therefore, a pivotal type prefers 0 to an even-odds (fifty-fifty) gamble between m and t under decreasing differences, but prefers the gamble under increasing differences.

With decreasing differences, the support of the type distribution is the darker shaded area in Figure 5 a) (with decreasing differences applied only to pivotal types, it is the whole shaded area). By reflecting over the line $t = -m$, we get the support for a type distribution with increasing differences. This is illustrated in Figure 5 b).

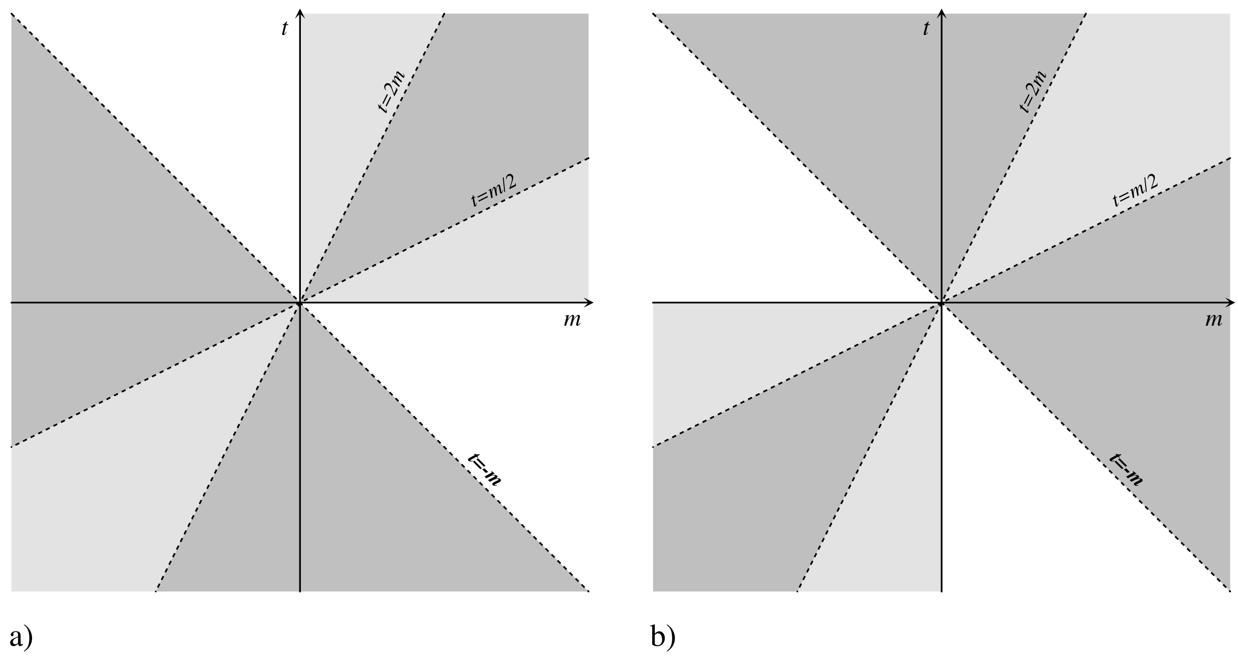


Figure 5: Types with Decreasing (left) and Increasing (right) Differences

Our introductory example satisfied decreasing differences.[11] There, we had $\left(u^1_A, u^2_A, u^3_A\right) = (1, 0, -2) = \left(u^1_C, u^2_C, u^3_C\right)$ and $\left(u^1_B, u^2_B, u^3_B\right) = (3, 2, 0)$. If we were to adjust these pay-

[9]The assumption that types who prefer M and T to zero are in the minority ensures that pivotal types are of consequence: at least some of them are needed to get a majority in favor of building the metro.

[10]One can see this directly from the definitions since, in the special case that $u^2_\theta = 0$, decreasing differences reduces to $u^1_\theta \leq -u^3_\theta$ and increasing differences to $u^1_\theta \geq -u^3_\theta$.

[11]If one has identified one type distribution that meets the assumptions of the game (and exhibits either decreasing or increasing differences), one can find more by raising the highest payoff and reducing the lowest payoff (by the same amount) for pivotal types, and raising the payoffs from metro and antiquity for the others (without changing the ordering and signs). This preserves the relative magnitude of differences and increases expected surplus from building, but does not affect voting majorities.

offs to obtain increasing differences for the pivotal types A and C, namely such that $(m_A, t_A) = (2, -1)$ and $(m_C, t_C) = (-1, 2)$, it is easy to see that, at any q, either A or C (if not both) would support digging and form a majority with B, since

$$qt_A + (1-q)\, m_A = -q + 2\,(1-q) = 2 - 3q \leq 0$$

only if $q \geq \frac{2}{3}$, which implies

$$qt_C + (1-q)\, m_C = 2q - (1-q) = 3q - 1 > 0.$$

Our main result regarding type distributions has two parts. First, a majority that opposes building the metro is always supported by some probability $q \in [0,1]$ that an antiquity is found, if the type distribution satisfies decreasing differences for pivotal types. Second, a majority always supports digging at any q, if the type distribution exhibits increasing differences for pivotal types.[12]

Proposition 2. *If the pivotal types satisfy decreasing differences, there exists a nonempty interval for q such that a majority opposes digging.*

Proof. We show that the votes of pivotal types (who rank 0 in the middle) are sufficient to achieve a majority that opposes digging, and that there exists a q such that both groups will in fact vote against it.

By assumption, types who rank 0 at the bottom ($M \succsim_\theta T \succsim_\theta 0$ and $T \succsim_\theta M \succsim_\theta 0$) are a minority. Those who rank 0 on top ($0 \succsim_\theta M \succsim_\theta T$ and $0 \succsim_\theta T \succsim_\theta M$) will certainly oppose digging, since that guarantees 0 and they prefer 0 to anYThing else. Thus, getting the remaining types, who rank 0 in the middle, to oppose digging is enough for a majority.

Any type θ opposes digging if $(1-q)\, m_\theta + qt_\theta \leq 0$. For someone with preference $M \succsim_\theta 0 \succsim_\theta T$, this is true if

$$q \geq \frac{m_\theta}{m_\theta - t_\theta} \equiv \underline{q}_\theta.$$

Similarly, someone with preference $T \succsim_\theta 0 \succsim_\theta M$ opposes digging if

$$q \leq -\frac{m_\theta}{t_\theta - m_\theta} \equiv \overline{q}_\theta.$$

With decreasing differences for pivotal types, $\underline{q}_\theta \leq 1/2$ for all θ such that $M \succsim_\theta 0 \succsim_\theta T$, since

$$\frac{m_\theta}{m_\theta - t_\theta} \leq -\frac{t_\theta}{m_\theta - t_\theta}$$

implies $q \leq 1 - q$. Similarly, $\overline{q}_\theta \geq 1/2$ for all θ such that $T \succsim_\theta 0 \succsim_\theta M$. Therefore,

$$0 \leq \max_{\theta \in M0T} \underline{q}_\theta \leq \frac{1}{2} \leq \min_{\theta \in T0M} \overline{q}_\theta \leq 1,$$

[12] These are statements about sufficiency; decreasing (respectively, increasing) differences is not necessary for them to hold. But they are about as weak as can be - any sufficient and necessary condition would have to restrict type densities in specific ways that would make the results more or less tautological.

which means there exists a q (in an interval that contains $\frac{1}{2}$) such that anybody who ranks 0 in the middle opposes digging. Given that less than half the population ranks 0 worse than both M and T, this is enough for a majority to oppose digging. ■

In a sense, decreasing differences "punishes" the introduction of a third alternative (finding the antiquity) through digging, which is the worst option for a significant part of the population (given our restrictions on the type distribution). The votes align against digging because those who would rank the antiquity last (the As in our initial example) want to shut this alternative out, even at relatively low q that make finding the antiquity unlikely. This eliminates intervals where those who like and dislike the antiquity might both support digging, since the latter require a lower q than the former are willing to accept. In consequence, the metro may not be built, even though it yields a net benefit and a majority desires it, and if the metro were abandoned later, a majority would consider that an even better outcome.

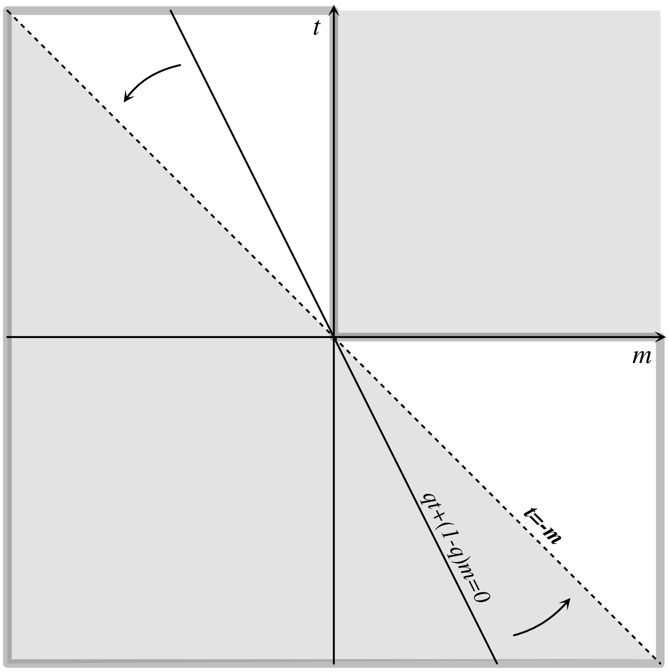


Figure 6: Inefficiency with Increasing Differences

Figure 6 illustrates how decreasing differences (so that the type support is the shaded area) in combination with the fact that a majority of types does not prefer both M and T to doing nothing (occupies the area inside the dark gray rim, where at least one of m and t is negative) brings about the possibility that a majority will oppose digging. As q approaches 1/2 (the slope of the indifference line approaches $q/(1-q) = -1$), more types fall below the line, until at $q = 1/2$ a majority is sure to vote against the project.

Proposition 3. *If the pivotal types satisfy increasing differences, a majority supports digging, regardless of q.*

Proof. Increasing differences for pivotal types implies

$$\frac{1}{2}m_\theta + \frac{1}{2}t_\theta \geq 0$$

for those who rank 0 in the middle, i.e. $M \succsim_\theta 0 \succsim_\theta T$ or $T \succsim_\theta 0 \succsim_\theta M$. Suppose, for the sake of contradiction, that a majority strictly opposes digging at some $q \in [0,1]$. Which types could comprise this majority? Type θ would strictly oppose digging only if

$$(1-q)\, m_\theta + qt_\theta < 0.$$

Given increasing differences for pivotal types, one of (i) $0 \succsim_\theta M$ and $0 \succsim_\theta T$, (ii) $M \succsim_\theta 0 \succsim_\theta T$ and $q > \frac{1}{2}$ or (iii) $T \succsim_\theta 0 \succsim_\theta M$ and $q < \frac{1}{2}$ has to be true then. Therefore, if $q < \frac{1}{2}$, types that satisfy (i) and (ii) must have a majority for there to be a majority that opposes digging. But then doing nothing is majority-preferred to T, which violates our payoff restrictions. If $q > \frac{1}{2}$, then types that satisfy (i) and (iii) must have a majority. Now, 0 is majority-preferred to M, another violation.

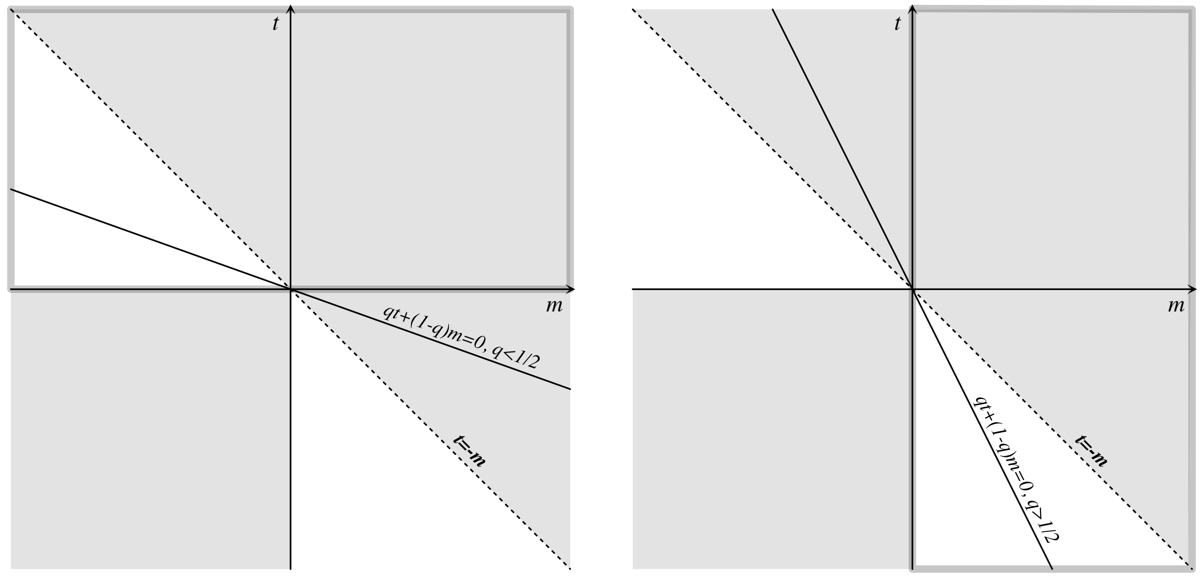
The only remaining possibility, $q = \frac{1}{2}$, is ruled out directly by increasing differences for pivotal types, since all who rank 0 in the middle would support digging, leaving only (i) to oppose it. But (i) cannot be a majority (again, because by assumption M and T are majority-preferred to doing nothing). Hence, there is no q that could produce a majority against digging and satisfy the basic assumptions of the game as well as increasing differences for pivotal types. ■

Increasing differences "rewards" the introduction of a third alternative through digging, which is the best option for part of the population. The votes align for digging because those who would rank the antiquity first (the Cs in our initial example) would like to enable this alternative, even at a relatively unfavorable (low) q. This creates an interval of qs where those who like and dislike the antiquity are both willing to support digging.

In Figure 7, we see how increasing differences ensure support for digging because of the assumptions that T and M are each majority-preferred to 0. If $q > 1/2$, the indifference line is relatively flat: its slope is $-(1-q)/q > -1$. Then the area shaded in the top panel lies completely above the indifference line. This area contains all types in the support who prefer T to 0 (i.e. for whom $t > 0$), which is a majority of the population. If $q < 1/2$, the indifference line is relatively steep with a slope smaller than -1. Then the area shaded in the right panel lies completely above the line. This area contains all types in the support who prefer M to 0 (i.e. for whom $m > 0$), and again this is a majority.

Decreasing differences for pivotal types is not necessary for a majority to oppose digging. Suppose, in violation, $(m_A, t_A) = (2, -1)$, while $(m_B, t_B) = (5, 6)$ and $(m_C, t_C) = (-4, 1)$, where each type has mass of one-third. There is a majority for T over M, and for M over 0. The average payoff from T (2) exceeds that from M (1), which exceeds 0. Despite the fact that decreasing differences does not hold for one pivotal type, a majority opposes digging at $q = \frac{3}{4}$, since

$$(1-q)\, m_A + qt_A = \frac{1}{4}(2) + \frac{3}{4}(-1) = -\frac{1}{4} \leq 0$$



a) b)

Figure 7: Efficiency with Increasing Differences

and

$$(1-q)\,m_C + qt_C = \frac{1}{4}(-4) + \frac{3}{4}(1) = -\frac{1}{4} \leq 0.$$

If we now increase t_C to 4, increasing differences is satisfied for the pivotal types. But then

$$(1-q)\,m_C + qt_C = \frac{1}{4}(-4) + \frac{3}{4}(4) = 2 \geq 0,$$

so we have a majority for digging, consistent with Proposition 3. In fact, for C to oppose digging, q would now have to be smaller than $\frac{2}{5}$, but no such q can induce A to oppose digging, i.e. solve

$$(1-q)\,(2) + q\,(-1) = 2 - 3q \leq 0.$$

There is a majority in support of digging at any q.

Two common classes of type distributions guarantee that the metro has majority support. We prove the following propositions in the appendix.

Proposition 4. *If the type distribution is symmetric about a point, a majority supports digging, regardless of q.*

In the case of point-symmetric distributions, it is impossible to divide the space with a line in such a way that the average voter (social planner) and the majority are on different sides, hence it is impossible for the majority to support an inefficient decision.

Proposition 5. *If valuations m and t are independently uniformly distributed (so that the type distribution is uniform on a rectangle), then a majority supports digging, regardless of q.*

A uniform distribution over a rectangle has the same basic property as a fully symmetric distribution - the majority and the average type always lie on the same side of any line. While it is tempting to generalize the previous proposition to uniform distributions on other (non-rectangular) shapes, this cannot be done. For example, if population preferences were uniformly distributed on a triangle, Proposition 5 would not hold in general.

6 Extensions

Our results in the previous section take advantage of the fact that the Roman metro game can be reduced to a decision between a choice (N) that leads to certain outcome (0) and a choice (Y) that leads to a lottery over two outcomes (M and T), which respect certain properties. (Namely that, in terms of majority preference and average payoff, T beats M, and M beats 0.) That some of these payoffs arise from further events (a vote between M and T, in case the antiquity is found) does not affect the proofs of Propositions 2 to 5. Hence, a similar inefficiency to that in the Roman metro problem arises in any game that can be reduced by backward induction to a choice between a certain outcome and a lottery over outcomes that satisfy the basic payoff restrictions, as well as decreasing differences.

Next we discuss a familiar job market situation where an inefficiency may arise.[13] Consider an economics department that considers flying out a job candidate for a macro position, who presents herself as both a macro and a labor economist. It is not clear where her primary interest lies, but if the flyout is scheduled, it will reveal the candidate's true field. After the candidate's field is revealed the faculty will vote on whether to hire her or to stick with the status quo candidate.

To some of the faculty (group A), it is important to hire a candidate who will exclusively work on macroeconomics; they would prefer to make an offer to the status quo candidate over a labor economist. Others (group B) are willing to make an offer to this candidate over the status quo candidate, regardless of whether she works in macro or labor. The remainder (group C) will vote for the candidate only if she is a labor economist, and otherwise prefer the status quo candidate.

In this case the uncertainty is over the type of the candidate. With probability q, the flyout would reveal that she is actually a labor economist. If q is neither very small nor very large, groups A and C may both refuse to fly her out, for different reasons. A worries that the candidate turns out to be a labor economist, and B and C then align to make her an offer. C worries that the candidate turns out to be a macroeconomist and is supported by A and B. Hence, the candidate may fail to get the flyout, even though, regardless of what the flyout reveals, a majority would vote to make her an offer!

The game is illustrated in the following Figure.

Here the utility from the status quo candidate is normalized to zero and the flyout is assumed to be costless in the sense that the status quo candidate yields the same payoff before and after the flyout.

Let us assume that $E(m) > 0$ and $E(t) > 0$ and that for at least half of the faculty,

$$t > 0$$

[13]This example was suggested by Jeff Ely.

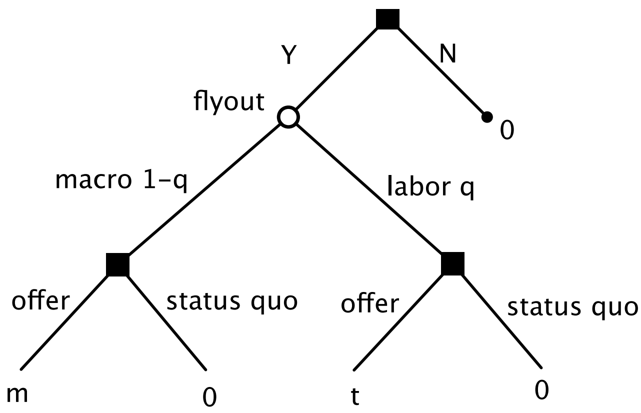


Figure 8: The job market game

and, for at least half of the faculty,

$$m > 0.$$

Given the above, regardless of what is revealed about the candidate, she will be hired over the status quo, and the social planner would also do the same. Nevertheless, as in the Roman metro game, an inefficient decision may be made not to fly the candidate out in the first place. Note that, given how majorities are allocated, and using backward induction, the choice to fly the candidate out amounts to playing a lottery, which gives t with probability q, and m with probability $1 - q$.

The inefficient decision not to fly the candidate out will be taken if for a majority $qt + (1 - q)m < 0$ because, by assumption, $qE(t) + (1 - q)E(m) > 0$. As with the Roman metro game, we can state the following proposition:

Proposition 6. *In the job market game, a majority opposes the flyout if and only if there exists a Condorcet cycle over the set of contingent plans that would be feasible with commitment.*

Proof. Here there are five possible contingent plans: do not fly the candidate out (N); fly the candidate out and hire if she is macro but do not hire if she is labor (YM_1L_0); fly the candidate out and hire if she is labor, but do not hire if she is macro (YM_0L_1); fly the candidate out and hire in any case (YM_1L_1); and finally, fly the candidate out, but do not hire in any case (YM_0L_0). By our assumptions, (YM_0L_0) and (N) produce the same payoffs. Also, (YM_1L_0) and (YM_0L_1) are majority-preferred to (N) because they are lotteries between (N) and something that is preferred by a majority to (N). (YM_1L_1) is majority-preferred to (YM_1L_0) and (YM_0L_1) because it shares one outcome with them and has another outcome which is majority-preferred. The only remaining comparisons are between (YM_0L_1) and (YM_1L_0), which - regardless of where the majority stands - cannot lead to a Condorcet cycle, and between (YM_0L_0) and (YM_1L_1). The only way the Condorcet cycle can arise is that (YM_0L_0) is majority-preferred to (YM_1L_1), which is exactly when the inefficient choice is made not to fly the candidate out. Hence, the majority opposes the flyout when there is a Condorcet cycle over contingent plans. ■

The job market game has an added dimension compared to the Roman metro game.

Candidates have some control over how they position themselves in the market. By identifying herself clearly as a macroeconomist or a labor economist, the candidate can remove any uncertainty, in which case the department makes the efficient choice and invites her. This rationalizes the advice job candidates often get from their advisors: avoid mixing fields.

The job market example illustrates that, while decision makers have no clear way out of the inefficiency - recall that pre-commitment to hire in a particular field, or not at all, only leads to a Condorcet cycle - candidates can fix the problem by revealing their type. The corresponding action in the Roman metro problem would be to locate antiquities prior to the decision over construction, so that the uncertainty is resolved. Unfortunately, this is not feasible in practice.

7 Conclusion

In the Roman metro problem, citizens decide whether to construct a metro line, which has a positive value for a majority, and which yields with some probability an option that a majority considers to be still better (finding a valuable antique). We presented an example in which a majority nevertheless prefers to do nothing instead (which is the worst possible outcome from an efficiency standpoint), then derived sufficient conditions both for an inefficient and an efficient majority choice (pro and contra construction). These are, respectively, a tendency to regard the worst outcome as too bad to be compensated by the best outcome (a property we call decreasing differences) and the opposite tendency (increasing differences). Under decreasing differences, the fact that digging and maybe finding an antiquity opens the door to a new alternative is a bad thing for many citizens. Under increasing differences, more possibilities are good. In this sense, one could look at the decision not to extend Rome's metro as a kind of field experiment, which reveals that citizens are preoccupied with the possibility of getting their least-preferred outcome.[14]

As a game form, the metro story fits other applications where "projects" unfold over time, with new possibilities arising as a result of initial actions. At the risk of oversimplifying complex realities, the logic of the Roman metro problem can be related to the civil uprisings in the Arab world, such as the current one in Syria, which is in danger of failing.[15] The population divides roughly into three camps: secular opposition, moderate Muslims who would like to establish a democracy, and religious fundamentalists, who want an Islamic state. Arguably, secularists would rather live under the current dictatorship than under Islamic law, while fundamentalists favor the dictatorship over democracy, which would grant liberties they view as offensive. Moderate Muslims prefer an Islamic state to a dictatorship.

Suppose a successful revolt requires that any two groups join forces. Ideally, it will lead to a popular vote on the country's future. But there is a risk that the situation would escalate and turn into a civil war between the secular and religious opposition. At the end

[14]This is not risk aversion, which would be refusal of a fair gamble with an equal upside and downside. Decreasing differences says, simply put, that the downside is regarded as large relative to the upside.

[15]This is a variation on an example that was first suggested to us by Steve Schmidt.

of the civil war, the dictator would be deposed, but it would be impossible to establish a political process that respects majorities. Instead, an Islamic state emerges. The lesson from the Roman metro problem is that the uprising may not succeed because of this risk. While moderate Muslims always support it, secularists participate only if the probability that civil war will break out is small, since they do not want to aid the creation of an Islamic state. Fundamentalists, on the other hand, will only act if the probability that an Islamic state results is high. They distrust the intentions of secularists and moderate Muslims to establish a democracy if the revolt succeeds and leads to a vote. Hence, if the probability of civil war is intermediate, secularists and fundamentalists may form an unlikely coalition to keep the dictator (whom it is in everyone's interest to remove) in power.[16]

Appendix

Proof of Proposition 4. If a distribution is symmetric about a point, that point is $(E(m_\theta), E(t_\theta))$: its coordinates are the population averages of m_θ and t_θ. According to our assumptions $(E(t_\theta) > E(m_\theta) > 0)$, this point has to lie in the positive quadrant.[17] Clearly, a hypothetical type that occupies the symmetry center will support digging, since this type prefers both antiquity and metro to 0. Every line that goes through the center of a symmetric distribution is a median (i.e. a line that divides the space into half-spaces containing equal mass).[18] In particular, the line with slope $-(1-q)/q$ passing through the center is a median. Since this line is parallel to, and lies above, the line with slope $-(1-q)/q$ passing through the origin, which is the indifference line, a majority values digging at least as much as the center. Since the center prefers digging, a majority must prefer digging, too.

Proof of Proposition 5. Consider a uniform distribution G on $[\underline{m}, \overline{m}] \times [\underline{t}, \overline{t}]$, a rectangle. The density is $g(m_\theta, t_\theta) = 1/\left((\overline{m} - \underline{m})\left(\overline{t} - \underline{t}\right)\right)$ for all θ. In general, an individual of type θ will oppose digging provided

$$(1-q)\, m_\theta + q t_\theta \leq 0,$$

i.e. when

$$m_\theta \leq -\frac{q}{1-q} t_\theta.$$

[16]Even if there were a way to commit to either a democracy or an Islamic state in the event that the dictator is overthrown, our first proposition shows that it would not solve the problem (if a majority opts not to revolt without commitment), since there is then necessarily a Condorcet cycle over the three alternatives (future democracy, future Islamic state, no revolt).

[17]Or, exactly at zero when the distribution is symmetric about zero, so that $E(t_\theta) = E(m_\theta) = 0$.

[18]More precisely (to allow for distributions whose support consists of disjoint sets), there exists, for every median, a parallel hyperplane (median) that passes through the center and divides the population in the same way.

Hence, a majority opposes digging if and only if

$$\int_{\underline{t}}^{\bar{t}} \left(\int_{\underline{m}}^{-\frac{q}{1-q}t_\theta} g\left(m_\theta, t_\theta\right) dm \right) dt \geq \frac{1}{2}.$$

Resolving the integral with uniform densities, a majority opposes digging only if (i) $\bar{t} + \underline{t} \leq 0 \leq \overline{m} + \underline{m}$ and

$$q \geq \frac{\overline{m} + \underline{m}}{\overline{m} + \underline{m} - \bar{t} - \underline{t}},$$

or (ii) $\overline{m} + \underline{m} \leq 0 \leq \bar{t} + \underline{t}$ and

$$q \leq \frac{\overline{m} + \underline{m}}{\overline{m} + \underline{m} - \bar{t} - \underline{t}},$$

or (iii) $\overline{m}+\underline{m}, \bar{t}+\underline{t} \leq 0$ (and any q). The conditions $\bar{t}+\underline{t} \leq 0 \leq \overline{m}+\underline{m}$ and $\overline{m}+\underline{m} \leq 0 \leq \bar{t}+\underline{t}$ imply that average payoffs rank M over 0 over T, respectively T over 0 over M.

For an individual of type θ, the expected surplus from digging is $(1-q)\, m_\theta + q t_\theta$. This is positive across individuals if

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \left((1-q)\, m_\theta + q t_\theta\right) g\left(m_\theta, t_\theta\right) dm \right) dt \geq 0.$$

In the case of the uniform distribution, the necessary condition for positive surplus from digging is

$$\int_{\underline{t}}^{\bar{t}} \left(\int_{\underline{m}}^{\overline{m}} \frac{(1-q)\, m_\theta + q t_\theta}{\left(\overline{m} - \underline{m}\right)\left(\bar{t} - \underline{t}\right)} dm \right) dt \geq 0,$$

which is true if and only if

$$(1-q)\left(\overline{m} + \underline{m}\right) + q\left(\bar{t} + \underline{t}\right) \geq 0.$$

This is consistent with two scenarios: (i) $\bar{t} + \underline{t} \leq \overline{m} + \underline{m}$ and

$$q \leq \frac{\overline{m} + \underline{m}}{\overline{m} + \underline{m} - \bar{t} - \underline{t}}$$

or (ii) $\overline{m} + \underline{m} < \bar{t} + \underline{t}$ and

$$q \geq \frac{\overline{m} + \underline{m}}{\overline{m} + \underline{m} - \bar{t} - \underline{t}}.$$

These directly contradict the conditions for a majority to oppose digging. (Except in the special case that $\overline{m} + \underline{m}, \bar{t} + \underline{t} \leq 0$, which is inconsistent with majorities for T and M against 0.) Therefore, uniformly distributed types on a rectangle guarantee support for digging.

References

Barbera, S., M. Maschler, and J. Shalev, "Voting for Voters: A Model of Electoral Evolution," *Games and Economic Behavior*, October 2001, *37* (1), 40–78.

Bernheim, B. D. and S. N. Slavov, "A Solution Concept for Majority Rule in Dynamic Settings," *The Review of Economic Studies*, 2009, *76* (1), 33–62.

Dutta, Bhaskar and Arunava Sen, "Implementing generalized Condorcet social choice functions via backward induction," *Social Choice and Welfare*, 1993, *10*, 149–160. 10.1007/BF00183343.

Jack, William and Roger Lagunoff, "Dynamic enfranchisement," *Journal of Public Economics*, May 2006, *90* (4-5), 551–572.

McKelvey, Richard D. and Richard G. Niemi, "A multistage game representation of sophisticated voting for binary procedures," *Journal of Economic Theory*, June 1978, *18* (1), 1–22.

Penn, Elizabeth Maggie, "A Model of Farsighted Voting," *American Journal of Political Science*, 2009, *53* (1), pp. 36–54.

Roberts, Kevin, "Condorcet cycles? A model of intertemporal voting," *Social Choice and Welfare*, October 2007, *29* (3), 383–404.

Srivastava, Sanjay and Michael A. Trick, "Sophisticated voting rules: the case of two tournaments," *Social Choice and Welfare*, 1996, *13* (3), 275–289.