A Mathematical Model of Dissent

Thomas Woodside

Philosophy of Dissent

April 30th 2023

Note: Paper is 16 pages but only 2998 words. The extra pages are a result of the use of graphs.

Revolution is fickle. Sustained effort from dissidents can lead to nothing until a sudden change occurs, as in Czechoslovakia. Similar dissident movements can develop different characteristics and lead to different outcomes for unclear reasons, as can be seen from the difference between the Maidan and the Belarusian dissent movement. In addition, those at the forefront of dissent are often punished severely, leading to a collective action problem. If one voice in a sea of many won't make much of a difference, why should anyone put their life at risk?

In this paper, I present a simple mathematical model of dissent. I argue that there are four broad types of dissenters, each of whom reacts differently to differing conditions of dissent. The model predicts that six factors are likely to influence the likelihood of dissent: assessments of the value of toppling the regime, assessments of the disvalue of being punished or arrested, the existence of benefits of dissent short of regime collapse, better information sharing about who is willing to dissent and when, and the number of people who feel obligated to dissent on principle even if there is little hope. All factors have been prominently targeted by successful dissidents, with different dissidents targeting different factors.

Types of Reasoning

Marginal egoistic reasoning

Marginal egoistic reasoning refers to reasoning about how taking a particular action would affect one's own wellbeing on the margin. It means asking the question, *if I took action x, how much benefit would I get, and how does that compare to the world where I did not take action x?* Marginal egoistic reasoning is common and often quite useful. Consider the following commonplace uses of marginal reasoning:

- Eating until you are full. The body provides a natural gauge of marginal utility to more food; if you are still hungry, it is positive, but once you are full, it may be negative.
- Buying something if it is below a certain price. If something is too expensive, the marginal benefit it brings is not worth the cost.
- Hiring an additional employee if you think they would cause you to add more revenue than you would spend on their salary and benefits. If they wouldn't bring in enough revenue to support their salary, they probably aren't worth hiring.

Economics tends to take marginal egoistic reasoning as a premise, and it is a common idea that people will consume or hire until their own marginal utility from that consumption or labor is equal to its marginal cost. It is furthermore assumed that people will make choices which maximize their utility, and by this economists mean the choices that will lead to the highest overall *marginal* utility.

However, marginal egoistic calculations can go wrong. A common way in which this can happen. In a prisoner's dilemma, two individuals (or groups) decide whether to cooperate with

one another, or defect. If they both cooperate, they share a large benefit. If either defects, they gain a small benefit that results from stealing the large benefit from the other player. If they both defect, they both lose the large benefit. The prisoner's dilemma is often used to model public goods, where some players can "freeload" at the expense of the others. The situation can be visualized as follows:

	Player 2: Cooperate	Player 2: Defect
Player 1: Cooperate	Player 1: large benefit Player 2: large benefit	Player 1: nothing Player 2: large and small benefit
Player 1: Defect	Player 1: large and small benefit Player 2: nothing	Player 1: small benefit Player 2: small benefit

If player 1 is a marginal thinker, they will think as follows: suppose that player 2 cooperates. I will be able to gain a large benefit if I cooperate, and a large and small benefit if I defect. So marginally, it makes more sense to defect. Similarly, if player 2 defects, I will gain nothing if I cooperate, and a small benefit if I defect. So it marginally makes more sense to defect. If the players are thinking from a marginal perspective, then they ought to both defect: but this would be worse for each of them than if they had both decided to cooperate. As such, marginal egoistic reasoning can be called *collectively self-defeating*: reasoning that does not succeed on its own terms when applied by a larger group of people.

There are two possible ways to escape from the suboptimal situation: altruistic thinking and total thinking.

Marginal utilitarian reasoning

One way to escape from the prisoner's dilemma is to act as a utilitarian: that is, to value the benefits and costs accruing to others the same as one's own benefits and costs.They are now unconcerned with *who* gets a benefit, only which benefits exist:

	Player 2: Cooperate	Player 2: Defect
Player 1: Cooperate	two large benefits	one large and one small benefit
Player 1: Defect	one large and one small benefit	two small benefits

For altruistic thinkers, it thus makes sense to cooperate in the prisoner's dilemma, and the optimal collective outcome is achieved.

Total reasoning

A second way to escape from the prisoner's dilemma is to assume that one ought to consider the total effect of taking actions like theirs, not just the effect of their own action. In the case of the prisoner's dilemma, one ought to consider the overall effect of people defecting, not just your own defection. If one is a total thinker, then one does not consider cases where players play different strategies, only where they play the same strategy.

|	Player 2: Cooperate	Player 2: Defect
Player 1: Cooperate	Player 1: large benefit Player 2: large benefit	Disallowed
Player 1: Defect	Disallowed	Player 1: small benefit Player 2: small benefit

If total thinking is applied, then cooperation is pursued by both players, solving the collective action problems. Total thinking was championed in different ways by Thomas Hobbes and Immanuel Kant. **Hobbes** believed that one ought to agree to a social contract requiring cooperation *if all others also agree to do so*.[1] **Kant** believed that one ought to take actions that one can will to be a universal law (in this case, cooperation) *in all cases*, regardless of the actions of others.[2]

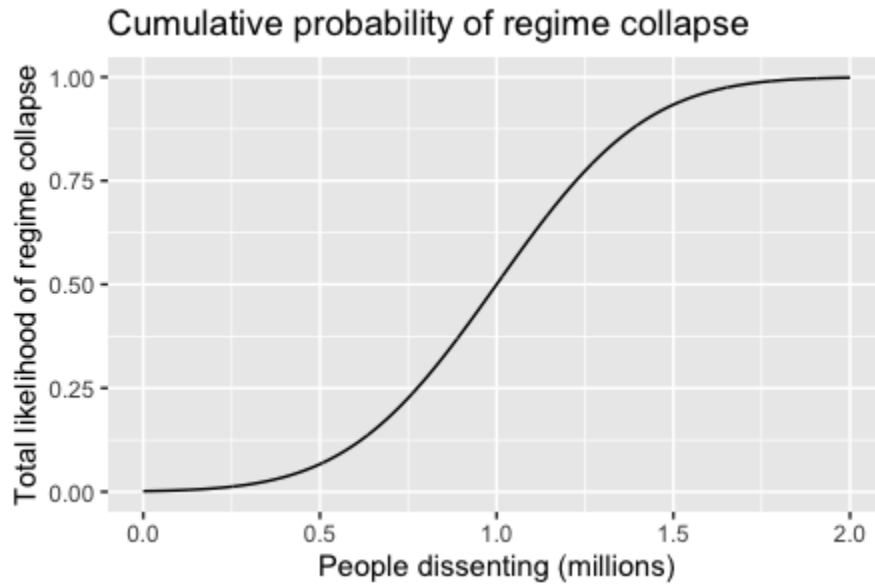## A mathematical model of dissent

A key aspect of dissent that differentiates it from the simple prisoner's dilemma is that nobody can be sure how many dissenters are needed to result in the collapse of the regime (or other desired change). It is clear an entire country's worth of citizens is enough to topple a regime, whereas a single citizen certainly isn't. In between, there is likely to be some level of uncertainty. This can be modeled starting with the assumption that people all wish to topple the regime, but it is unclear how many would be needed to do so:

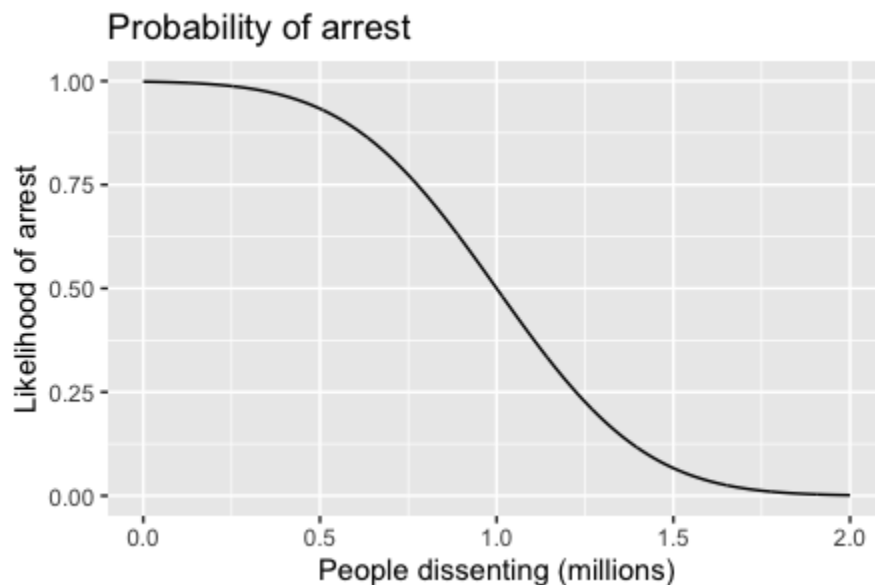1 Thomas Hobbes, *Leviathan*, 1651.
2 Immanuel Kant, *Groundwork for the Metaphysics of Morals*, 1785.

Number of dissenters needed to trigger regime collapse
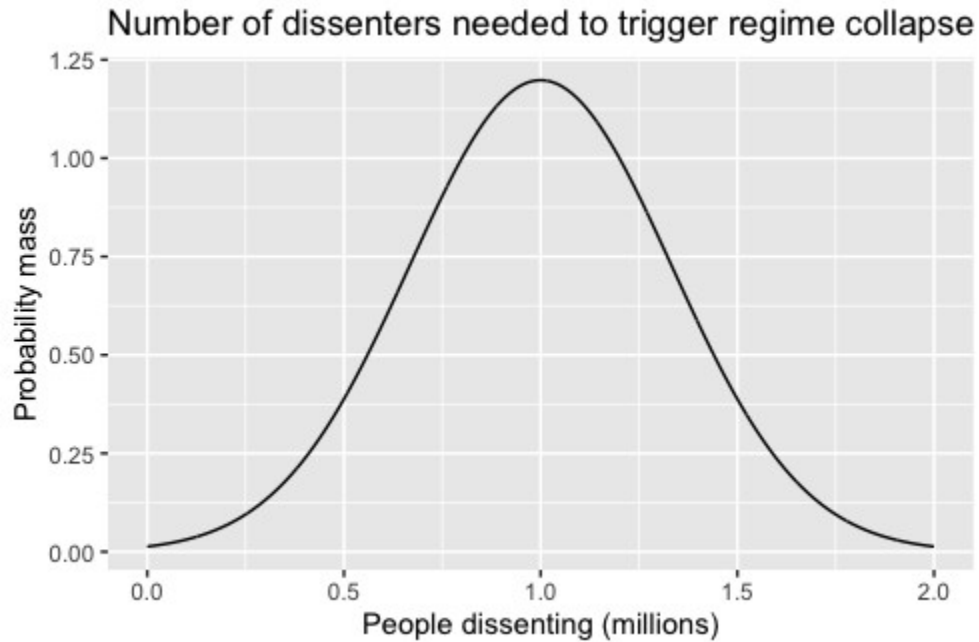
The graph above gives a hypothetical estimate of the number of dissenters needed to topple the regime. The mean estimate is that it would take 1 million people dissenting in order to cause the collapse of the regime. However, as can be seen from the graph, there is some possibility that the change could be triggered with far fewer people, or that it would need far more. To visualize this better, we can look at the *cumulative* probability of the regime falling. The graph below represents the probability the change will occur given that *at least* that many people dissent:

Cumulative probability of regime collapse

A line chart titled "Cumulative probability of regime collapse." The x-axis is labeled "People dissenting (millions)" ranging from 0.0 to 2.0. The y-axis is labeled "Total likelihood of regime collapse" ranging from 0.00 to 1.00. The curve follows an S-shaped (sigmoidal) increase, starting near 0.00, passing through about 0.50 at 1.0 million, and leveling off near 1.00 after about 1.75 million.

It is further important to model the fact that *until the regime falls, there are major costs to dissenters.* We can suppose that dissenters who dissent before the collapse of the regime is guaranteed will be arrested at least once:

Probability of arrest

A line chart titled "Probability of arrest." The x-axis is labeled "People dissenting (millions)" ranging from 0.0 to 2.0. The y-axis is labeled "Likelihood of arrest" ranging from 0.00 to 1.00. The curve follows a reverse S-shaped (sigmoidal) decrease, starting near 1.00, passing through about 0.50 at 1.0 million, and leveling off near 0.00 after about 1.75 million.

In this model, the first people to dissent will almost certainly be arrested, whereas after two million people have already dissented, arrest is nearly impossible as the regime is likely to

have already collapsed. In addition to the probability, we can also plot the *cumulative number* of

arrests given a certain number of dissenters. The graph below shows how many people we expect

to have *ever* been arrested, including those who might later be released if the regime falls:

Expected total number of arrests
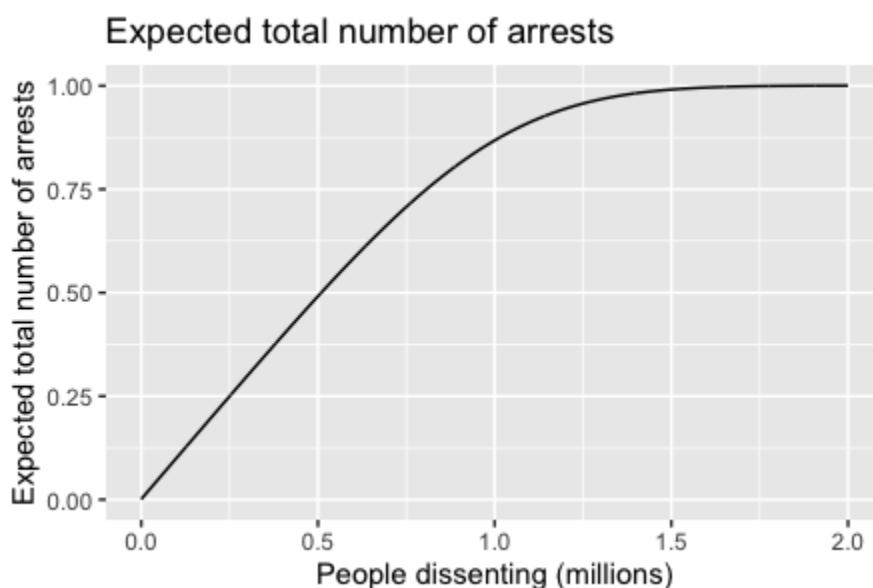
The model now has values for two important quantities: how likely citizens are to receive

the benefits of dissent (the collapse of the regime) and suffer the costs (arrest). We can now

begin to establish *trade-offs* between costs and benefits. Suppose that all two million citizens in

the country would prefer the regime to end, and furthermore, would be willing to be personally

arrested twice in exchange for a *guarantee* that the regime would end. In this situation, we can

value the collapse of the regime as worth at least four million arrests.

We can now calculate the *expected utility of dissent* by multiplying the probability of

regime collapse by 2 and subtracting the total number of expected arrests.
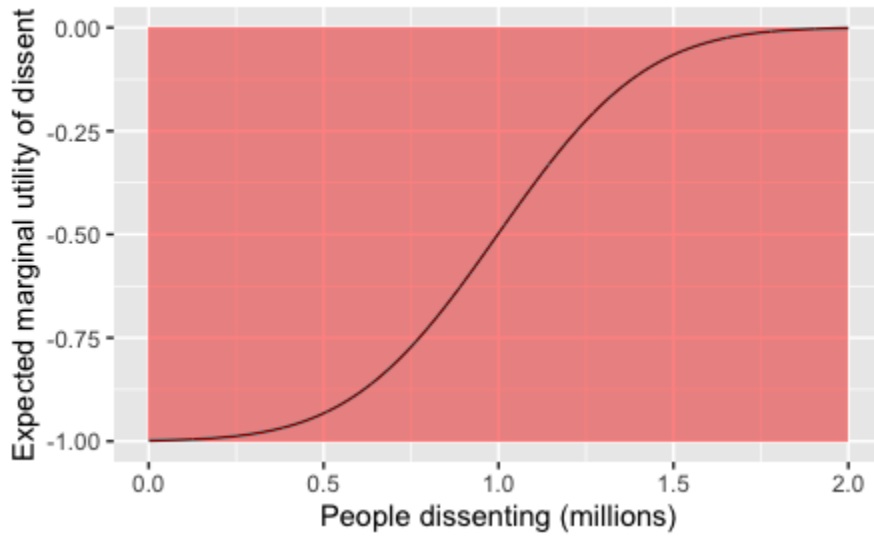
Expected total utility of dissent
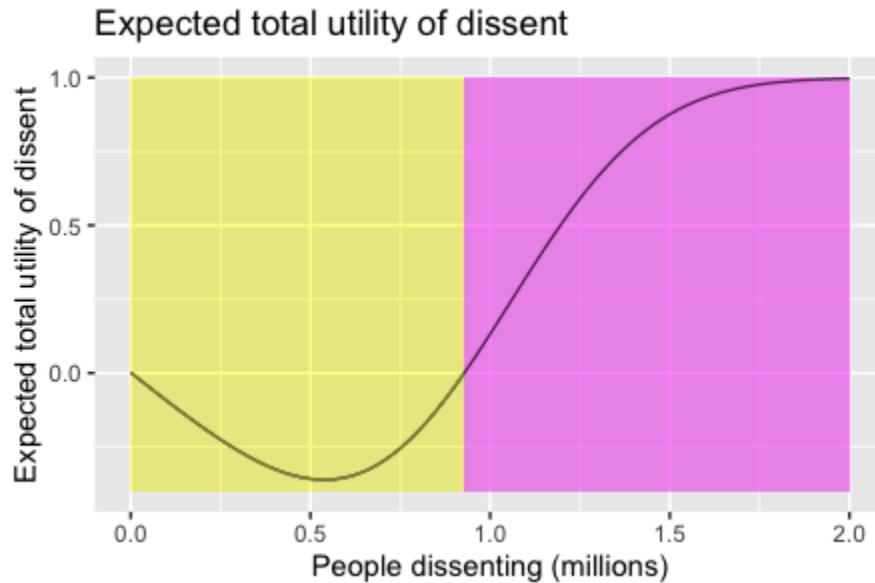
As can be seen from the graph, the initial stages of dissent have negative expected utility (yellow region). People are arrested, but it is highly unlikely that anything comes from it. However, once more people begin to dissent, the expected utility becomes positive and continues increasing until the end of the regime is virtually guaranteed (purple region).

The model assumes that every citizen would be willing to be arrested twice for a guarantee that the regime would end. Of course, no citizen can have such a guarantee. Citizens thinking marginally can only assess the (extremely small) probability that they will be the tipping point that personally leads to the end of the regime, and then subtract the cost of their own arrest.

Both egoist and utilitarian marginal thinkers will consider the expected cost of their own arrests, as well as the probability they will be the tipping point. However, egoists only consider the benefit of the regime ending to them personally, whereas utilitarians consider the benefits to all. Their marginal utilities are pictured below:
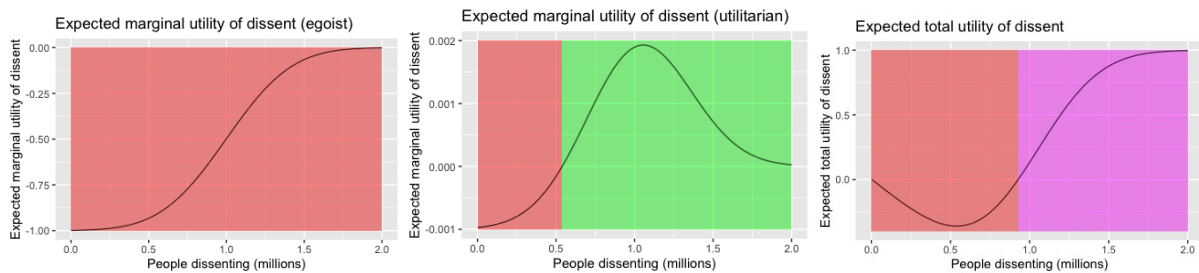
Expected marginal utility of dissent (egoist)

Expected marginal utility of dissent (utilitarian)
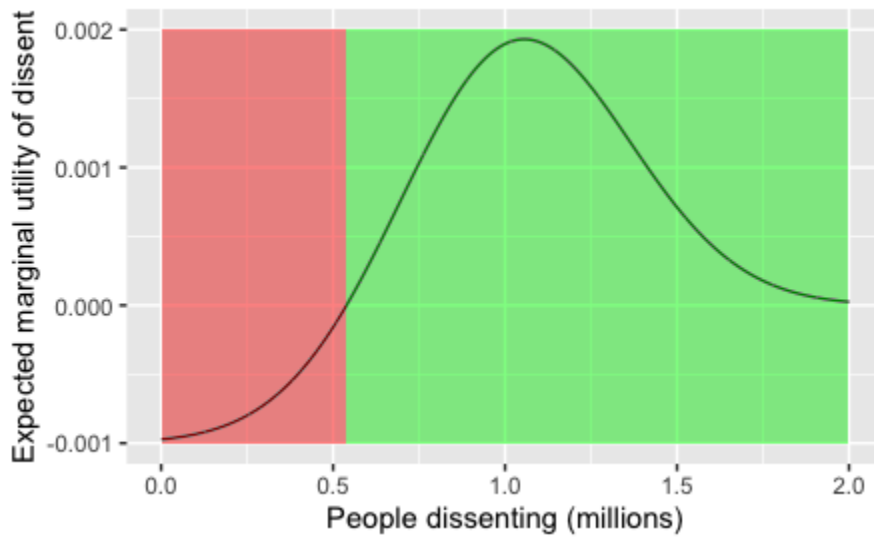
Different dissenting styles

Now that we have a model, we can evaluate how it predicts each sort of thinker will act. The graphs are presented again for ease of reference.

Egoistic marginal thinkers know that the probability that they themselves would lead to the end of the regime is far too low to be worth the much higher probability of arrest from a purely egoistic perspective. As such they **never dissent.**

Utilitarian marginal thinkers will dissent, but only when they have a good shot of personally being the "straw that breaks the camel's back." Otherwise, the benefits to everyone (a very small probability of regime collapse) don't outweigh the costs (a single additional arrest). These thinkers won't dissent until after many others are already dissenting (in this case, ~500,000). To summarize, they **dissent if enough people are already dissenting (green box).**

Hobbesian total thinkers consider the likelihood that *they will be part of a coalition* that topples the regime (regardless of whether it would have happened anyway without them). They multiply this by the benefit everyone would accrue if the regime falls, and weigh this against the harm that *everyone* would accrue from being arrested. Hobbesian total thinkers will also only dissent if they think they have a good chance of toppling the regime, but rather than being reactive, they are anticipatory: if they believe that there are enough (in this case, ~900,000) others who would agree to dissent, they will take to the streets, even if the others are not yet on them. To summarize, they **dissent if they anticipate enough others will also agree to dissent (purple box).**

Kantian total thinkers always dissent as long as dissent would be positive with any sufficiently large number of dissenters (in other words, if there is a purple region at all). Kantian total thinkers behave as though they are willing universal law for all other rational beings, and as

such do not consider whether other beings *really will* obey that law. I call them Kantians not because I expect they have read Kant, but because Kant is the most noted exemplar of this kind of reasoning. Kantian total thinkers will **always dissent as long as there is some coalition which *could* produce a lead to collapse (i.e. as long as the purple box exists).**

Application of the model to dissent in central and eastern Europe

While the model is quite simple, it does make predictions about which factors are likely to lead to an increase in willingness to dissent. All of the factors were in fact promoted by different dissident thinkers covered in this course.

For everyone, dissent becomes more likely when people have **increased assessments of the value of toppling the regime.** Havel's campaign on behalf of the Plastic People of the Universe is a good example of this, because it laid bare the fact that the regime was stopping people from living their lives, even if they did not oppose the government. As Havel wrote, "people in different milieus very quickly began to understand that a threat to the freedom of these young people was a threat to the freedom of us all, and that a strong defense was all the more necessary because everything was against them."[3] This kind of awareness-raising increases people's perceptions of how negative the regime is, and conversely how positive it would be for it to fall.

For everyone, dissent becomes more likely when people have **reduced assessments of the disvalue of punishment or arrest.** In Belarus, women who were imprisoned spoke of the solidarity that came about between women. Hanna Komar wrote:

---

[3] Václav Havel and Karel Hvížďala, *Disturbing the Peace*, trans. Paul Wilson (New York: Alfred A. Knopf, 1990), 130.

"When I asked a fellow inmate what helped her get through the prison experience recently, her response was unequivocal: 'knowing that it was going to end… and the other girls, of course'."[4]

This sort of solidarity, which extends not just to women in prison but also after they are released, certainly has the effect of reducing fear of imprisonment.

For marginal utilitarians, dissent becomes more likely when there are **benefits to dissent short of total regime collapse**. Marginal utilitarians need to feel as if there is some value to their individual contribution in order to participate in dissent. The idea of the parallel polis supports exactly this; it provided something for dissenters to *build*, rather than an illusory possibility of allowing them to take something down:

> "I suggest that we join forces in creating, slowly but surely, parallel structures that are capable, to a limited degree at least, of supplementing the generally beneficial and necessary functions that are missing in the existing structures, and where possible, to use those existing structures, to humanise them."[5]

In the absence of a parallel polis, marginal utilitarians will tend to become revisionists. Michnik was speaking of marginal thinkers when he said, "faith in one's ability to exert influence on the fate of society is an absolute prerequisite for political activity."[6] It is important for dissidents to remember that marginal utilitarians can be coopted by the regime, which may appear to offer a greater possibility of marginal benefit. Michnik wrote of revisionists:

---

4 Hanna Komar, "Sisterhood behind Bars," August 9, 2021, https://www.eurozine.com/sisterhood-behind-bars/.
5 Václav Benda, "The Parallel 'Polis,'" in *Civic Freedom in Central Europe: Voices from Czechoslovakia*, ed. H. Gordon Skilling and Paul Wilson (MacMillan, 1978), 36.
6 Adam Michnik, "A New Evolutionism," in *Letters From Prison And Other Essays*, trans. Maya Latynski (Berkeley: University of California Press, 1976), 137.

"Many of those who joined the PUWP defended their decision in the following manner: 'This way I will be able to serve the cause of Polish democracy, because in this way alone I will be able to lend effective support to the Polish Dubek when he appears.' So far, this service to the cause of democracy has amounted to service to the totalitarian powers."[7]

For Hobbesian total thinkers, dissent becomes more likely when there is **better information about who else is willing to dissent and when**. Since Hobbesian total thinkers are only willing to dissent when enough others are also willing to, they need to form coalitions with others thinking of dissenting before they themselves do so. The parallel polis is also a good example of this. But perhaps an even more clear example is Charter 77, where hundreds of people publicly declared themselves as being willing to dissent.[8] Vaclav Havel said that it was the campaign for the Plastic People that led to the information sharing necessary to form Charter 77, saying that "the main opposition circles, hitherto isolated from each other, came together informally."[9]

Dissent also becomes more likely if there are **more Kantian total thinkers**. This is especially important if the mere sharing of information itself is subject to restrictions, because without Kantian total thinkers willing to gather information and set up parallel structures, both marginal utilitarians and Hobbesian total thinkers would be unwilling to dissent. Many of the thinkers in this course did so by arguing that individuals were responsible for the harms of the regime if they did nothing, thereby creating a categorical reason to dissent. For example, Havel's *Power of the Powerless* is a critique of marginal reasoning:

---

7 Michnik, 139.
8 "Declaration of Charter '77," January 1, 1977.
9 Havel and Hvíždala, *Disturbing the Peace*, 131.

"[people] need not accept the lie. It is enough for them to have accepted their life with it and in it. For by this very fact, individuals confirm the system, fulfill the system, make the system, are the system."[10]

Jan Patočka meant something similar when he said that sacrifice for "no existing particular" allows us to win our humanity.[11] The idea of sacrifice for nothing would not appeal to a marginal thinker, and Patočka is asking for a rejection of that sort of marginal thinking. Finally Karel Kosík in Reason and Conscience decried the idea, held by marginal thinkers, that it is against reason to favor "real" over "illusory" probabilities.[12] Although Kosík did not frame it in the way discussed in this paper, one way of interpreting the comments is that he believes people may sometimes need to take actions even if they see no marginal benefit to them.

Concluding thoughts

This paper introduced a model of dissent and applied it to people who reason in different ways. The model predicts many of the strategies used by dissidents in central and eastern Europe to catalyze dissent. The framework suggests that a (potentially small) contingent of Kantian total thinkers may be necessary to catalyze any form of dissent at all. While all of society can aid in a revolution, only Kantian total thinkers can start one.

10 Václav Havel, *The Power of the Powerless*, ed. John Keane, trans. Paul Wilson (Armonk, New York: M. E. Sharpe, Inc, 1979), 31.
11 Jan Patočka, "The Dangers of Technicization in Science According to E. Husserl and the Essence of Technology as Danger According to M. Heidegger," in *Jan Patočka, Philosophy and Selected Writings*, ed. Erazim Kohák (Chicago and London: The University of Chicago Press, 1989), 339.
12 Karel Kosík, "Reason and Conscience," in *The Crisis of Modernity: Essays and Observations from the 1968 Era*, ed. James H. Satterwhite, trans. Julianne Clarke (Lanham, MD: Rowman & Littlefield, 1995), 14.