A Computational Model of Workers Protest

Jae-Woo Kim* and Robert A. Hanneman**

February 10, 2009

* Department of Sociology, University of California, Riverside, CA 92507, USA
** Department of Sociology, University of California, Riverside, CA 92507, USA

Abstract

This paper presents an agent-based model of workers protest. Workers have different degrees of grievance depending on the difference between their wage and the local average. Workers protest with probabilities proportional to grievance, but are inhibited by the risk of being arrested which is determined by how many probable rebels compared to coercive agents exist in their local areas. We explore the effect of perception of similarity on the dynamics of collective actions. If workers are surrounded by more in-group members, more risk-taking; if surrounded by more out-group members, more risk-averse. We examine how individual interest and tag-mediated group membership identity interact with each other on patterns of workers protest: its frequency, inter-occurrence time, strength, and durability. The simulation experiment results indicate that rhythms of outbursts fit into recurrent events with negative time dependence rather than independent events. Group identities do not influence seriously how often local uprisings emerge. Both how strong global protest is and how long it is sustained are significantly sensitive to tag-mediated risk aversion.

Keywords: Agent-based Simulation; Workers Protest; Tag-based Group Identity; Trust; Netlogo

Introduction

1.1 It is the marked increase in wage inequality, and major layoffs due to corporate restructuring brought on by a financial crisis that triggered a major outbreak of strikes and local protests in South Korea in 1996-97. Although their wave was severe in industrial cities, it never reached its full potential as far as solidarity between Korean workers and immigrant workers primarily from Southeast Asia who were more underpaid with the higher risk of getting fired than native workers. Ultimately, coercive efforts by employers and the government were able to prevent the emergence of a synchronized worker action that spans both spatial and ethnic diversity in the working class.

1.2 In this paper, we extend a model of protest mobilization proposed by Epstein (2002) to examine effects of perceived "in" and "out" group membership identity on the shape of protest events. In the model, workers are motivated to protest by grievances based on wage inequality, but it is inhibited by the risk that coercive agents might arrest them. Basically, variations in grievance and coercion give rise to dynamic patterns of protest activity. To this model, we add the distinction between two types of workers (i.e. natives and guests) who can identify others in their local neighborhood by observable "tags" of physical appearance or support of others from their own social group, and that this modifies their risk calculus. In a series of simulation experiments, we explore the consequences of this "who are trustworthy and who not" for the pattern of protest waves. In this way, we propose an agent-based model of worker protest to investigate how individual interest and group membership identity work together in its dynamics.

Theoretical Frame

2.1 The rational choice theory axiomatically states that individuals behave in a rational manner given perfect information to maximize calculable utility in terms of cost-benefit. This individual-based utilitarian model has been widely implemented to explain collective behavior. Simply saying, agents will participate in collective behavior if perceived benefit – perceived cost > threshold; otherwise, not.

2.2 The dynamics of collective behavior among heterogeneous agents is not the same with its dynamics among homogeneous ones. There are two distinguishable approaches to theorizing and modeling it. First, studies on the critical mass (Granovetter, 1978; Marwell and Oliver, 1993; Brichoux and Johnson, 2002) highlight roles of a small group of people who are different from the others in terms of interest, resource, and/or commitment. Second, given the absence of identities in the rational choice theory, we are more interested in the issue of trust based on group membership identities among heterogeneous agents and its impact on individual preference.

2.3 Group identities in our study are different from either individual identities or collective identities. Individual identities are meanings individuals assign to themselves. Such identities can be absorbed into utilities in a broad sense to be defined as a function of individual choice, participation or not. A social constructivist approach to collective identities – which are more transient and more emergent rather than rooted in socio-cultural categories (Rohlinger and Snow, 2003) – is not our focus. Our study is more impelled by the categorization process (Tajfel, 1974) and "tagging" (Holland, 1996). Making a distinction between "us" and "them" based on observable cultural markers (e.g. ethnic markers[1]) is a fundamental mechanism in group processes.

2.4 Human agents rely on cognitive cues to identify themselves and others, and they are more likely in a large size of gathering with strangers (i.e. workers from other workplaces in our model). Labeling predicated on similarity perception among heterogeneous agents interplays with individual interest in the dynamics of collective actions. Particularly in a protest, people in front of coercive agents take into account whether they are surrounded by an enough number of trustworthy people as well as the net risk determined by the difference between perceived benefit and perceived cost.

Agent-based Modeling of Collective Behavior

3.1 Many examples of "emergence" have insights for collective behavior: how a flock of birds in the sky keep their movements so orderly without the leader?; how clapping (or noise) at some corners of a large concert (or lecture) diffuses, amplifies, and disappear?; how crickets synchronize their chirping to make a series of rhythms with a variety of strength (Watts, 2003)?; and how a large size of flip card section can be coordinated even though participants do not have global vision?

3.2 Imagine first that a huge number of hooligans sit together in a large stadium to watch the match between LA Dodgers and NY Yankees. Whoever wins the game is supposed to go to the upcoming international league. Unfortunately, cops are asked to have seats here and there with

[1] One might ask how the theory of split labor market (Bonacich, 1972) is related to our study. We are concerned with the interplay between class and race, but underpaid workers are more likely to participate in collective behavior in our model. The opposite is the case with that theory, according to which native workers are forced to compete with guest workers who are more likely to remain as bystanders or strikebreakers.

hooligans to arrest and send them to the jail in order to preclude violent conflicts if they make an audience wave which might be able to stir up antagonism. Both hooligans and cops can move out to another vacant seat if there is any within a certain radius. Hooligans have different degree of grievance depending on the score difference between two teams. The more likely for hooligans to make a wave, since the more nearby comrades and less cops surround them within the same radius, the less perceived risk of being arrested. Here is another scenario: hooligans might want to count upon cognitive cues such as physical appearance and accent to distinguish one group of audience from the other? What if they trust perceived in-group members more than out-group members?

3.3 You could imagine audience waves diffuse across the space through chain reactions. What do they look like in terms of frequency, strength, and durability? Are there any significant differences between the first scenario and the second one in those patterns? Agent-based modeling from the bottom-up perspective is fundamentally useful to tackle these questions. We extend an agent-based model of civil violence (Epstein, 2002) to investigate how individual interest and tag-mediated group identity interplay in a different context – workers protest.

The Model[2]

4.1 The latest version of Netlogo (Wilensky, 1999) is equipped with Model I in Epstein (2002) in her library which we use as a starting point. Workers are located on a grid with some open spaces instead of civilians. In his original model, both perceived hardship (H) and perceived regime legitimacy (L) determine grievance: G=H(1-L), H~U(0,1), but L is fixed as a certain value. Instead of H and L, we define each worker's grievance as a feeling of relative deprivation which motivates to consider joining a protest:

$$G = 2 \times \left| \frac{1}{1+e^{-D}} - 0.5 \right| \quad (1)$$

where D is ego's wage – local average within a certain radius. G has 0 as the minimum when D is 0 or bigger. Its maximum is 1. The exponential function (Figure 1) reflects that grievance changes more sensitively, less difference, but less, bigger difference. Wage is any random number from $N(0.5, 0.167^2)$ multiplied by wage dispersion (WD, hereafter). It is assumed that wage does not change throughout 300 steps. If WD is 1, Wage~$N(0.5, 0.167^2)$; If WD is 2, ~$N(1, 0.334^2)$ since $E(aX)=aE(X)$ and $Var(aX)=a^2Var(X)$; and so on.

[2] Our model applet is available at http://student.ucr.edu/~jkim081/simulation.htm

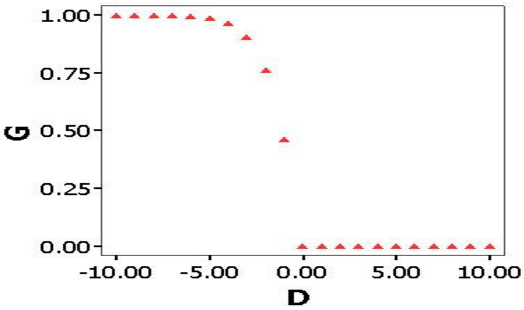

Figure 1. Plotting Equation 1

4.2 When WD is 3, we could reproduce almost the same initial value of the average grievance compared to Epstein's model at the setting that perceived hardship and perceived legitimacy are U(0,1) and 0.82, respectively. However, there is a significant difference in the dispersion of grievance between two models. It seems that the overall distribution of grievance in our model (Left) is more realistic than his (Right) as is shown in Figure 2.

Figure 2. Comparison of Two Distributions of Grievance

4.3 Each worker chooses to protest or not, which is contingent on a rational calculation of their net risk. Net risk, in turn, is determined by an individual-level risk-aversion and an estimated probability of being arrested. In other words, Net Risk (N) equals with Risk Aversion (R) multiplied by Estimated Arrest Probability (P) as is in Epstein's model. But, we used $R \sim N(0.5, 0.167^2)$ instead of $R \sim U(0,1)$ in his model. In this way, the distribution of risk aversion is exogenously fixed. Here, $P=1-\exp[-k(C/A)_{vision=6}]$. C/A is the number of cops over the number of workers who have grievance bigger than risk plus threshold within a certain radius, the 6-patch neighborhood here. Suffice to say that workers know only what is going on in their immediate surroundings (i.e. local vision without leaders). Following Wilensky (1999), we set k as 2.3 in "startup" to ensure a reasonable value when there is only one cop and one agent within a particular field of vision and also take the "floor" of C/A – the largest integer less than or equal to C/A. Overall, the more (or less) active workers around me, the lower (or higher) perceived arrest probability.

4.4 At the beginning of each round of simulation with updating asynchronous, both workers who are not arrested and cops move to one of vacant cells within a 6-patch radius[3]. And then, workers become active if the difference between their grievance and net risk falls above a threshold (0.1). Let us suppose that companies ask cops (i.e. coercive agents) to arrest active workers. If they are arrested, they have to get fired temporarily and back to fellows after a certain period of time. For this reason, we set Layoff-term instead of Jail-term in Epstein (2002)[4]. It is a random real number between 0 (i.e. agents are not laid off) and 30 (i.e. maximum). This random number approach can be justifiable if employment flexibility differs from one workplace to another. For example, when one active worker is laid-off and 13 assigned to her, she can get back to the population after 13 steps with the same wage, tag value, identity comparator, and risk-aversion, as is for risk aversion and perceived hardship in Epstein (2002).

4.5 We explore effects of identity heterogeneity on the shape of protest events. It is modeled as a continuous variable – which could be understood as an agent's cognition and scaling of overall similarity to each neighbor. Whether the difference between an ego and her alter is labeled as "othering" is determined by the level of tolerance of each actor to "tag" difference. The tag value of each agent can be any real number between 0 and 1, following Riolo et al. (2001). Different values are assigned to workers at the beginning, and let me assume that they do not change throughout 300 steps. Labeling depends on two factors: | ego's value – alter's value | and tolerance for dissimilarity (T) ~N(0.5, 0.167^2). For instance, if 0.6(ego A), 0.2(alter B), and T_A is 0.5, | 0.6 – 0.2| < 0.5. Hence, A categorizes B "us." For another ego C with 0.9 as the tag value and the same level of comparator (T_C=0.5), B is recognized as "them" since | 0.9 – 0.2 | > 0.5. In general, if the absolute value of the tag difference is less than or equal to T, in-group; otherwise, out-group.

4.6 First, by default when there is no tag-mediated group identity, if | grievance – perceived risk | > threshold, join the protest; otherwise, remain as a bystander. Second, if tag-mediated distinction between us and them is manipulated, workers become active when the number of perceived in-group neighbors is the half or more than the number of out-group members, besides the first

---

3 Epstein (2002) shares the assumption of local vision with Brichoux and Johnson (2002), but agents can move around within their local neighborhood. This makes his model more realistic and reasonable.

4 It should be noticed that Epstein (2002) could observe punctuated equilibrium because he set the maximum jail term relatively high (e.g. 30). If the maximum is small enough (e.g. 10), we can see some agents always active during the rest of steps once a strong local protest occurs. In this way, his model is very sensitive to the initial value of the maximum jail-term. He was not serious about this issue, conjecturing that increasing the jail term would 'flatten' the distribution of waiting time between protests and raise its mean (Ibid: 7246). Our model does not show such sensitivity.

condition (i.e. G–N > 0.1). Whether a random integer number generated from the uniform distribution is bigger than a given level of tag-mediated distinction parameter determines which one occurs between the first process and the second one (e.g. given 20% of tag-mediated labeling, if 35, the first process; and if 16, the second). Agents' decision-making is purely rational at 0% tag-mediated perception, but tag-based rational at 100% (i.e. tag-based risk perception + Net Risk calculation).

Experimental Design and Variables

5.1 In the factorial experimental design in Table 1, the effect of tag-mediated labeling, the effect of wage dispersion, and their interaction effect are investigated when holding all other variables constant. Given 100 runs with 300 steps at each of different settings, the number of independent cases is 1,200 in total.

Table 1. Experimental Design

Experiment ID	Tag-mediated Labeling	Wage Dispersion
1	0	2
2	0	4
3	20	2
4	20	4
5	40	2
6	40	4
7	60	2
8	60	4
9	80	2
10	80	4
11	100	2
12	100	4

Notes: Controlled variables are: 1) Initial density of cops (4%, 18 in total), 2) Initial density of workers (70%, 309 in total); 3) Vision (6 patches); 4) Threshold (0.1); 5) Maximum layoff-term (30); and 6) Movement (Yes).

5.2 Patterns of protest cycles are of great interest in political sociology and the study of social movements (Olzak, 1989; Oliver and Myers, 2002). Tilly (1978) suggests that individual events can be usefully characterized as having more or less severity (i.e. how many agents are involved at the peak), duration (i.e. how long does the event last), and volume (i.e. the sum of the duration of agent's involvements during an event). The distributions of protest rhythms may be exponential or scale-free. They may display regular cycles, varying degrees of autocorrelation, or varying degrees of chaotic/ regular patterning.

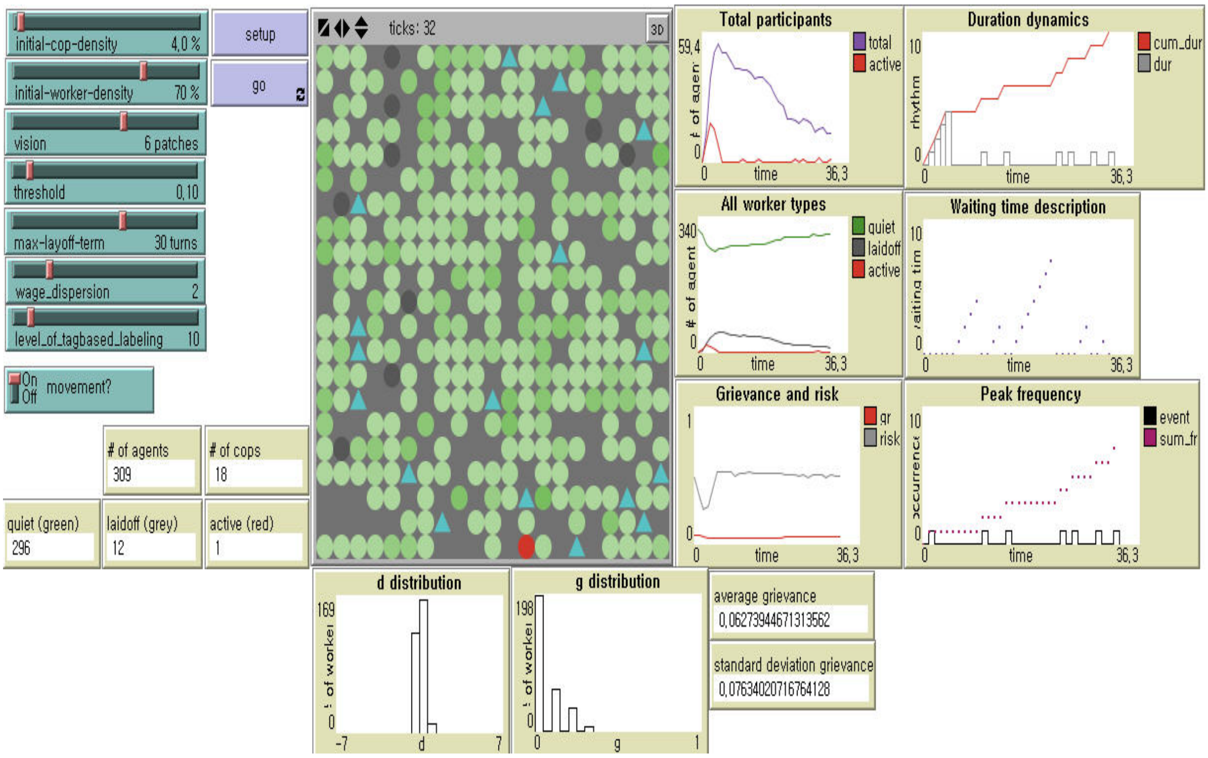

Figure 3. Model Interface
Notes: Cops (Triangle); Workers (Circle) – Rebel (Red), Laid-off (Black), Bystanders (Green, but its tone reflects the strength of grievance)

5.3 An event is regarded as a protest when at least one worker is active without being arrested. The event ends when the number of active workers falls to zero immediately after all rebels are laid-off. The following five outcome variables are supposed to be measured; 1) peak frequency is defined by how often there are cycles of protest. In Figure 3, for example, we see seven protest events through the first 32 time-steps; 2) waiting time is measured by how long a peaceful period will last between events. The waiting time list is recorded as [0 4 2 7 0 2 1 0] in the output text file, but waiting time from step 0 to the first event is regarded as censored to be excluded. In the similar manner, the last item will be deleted if the last event is still going on to the end as is in Figure 3. This is how [4 2 7 0 2 1] is recorded as the final waiting time list for statistical analysis; 3) the height of the maximum peak as the acme of activation is the maximum number of active workers at any point during the run. It is 18 up to 32 steps which is the same with the number of active workers in the first peak; 4) the average strength of the protest is measured by the number of active workers per step, that is, the total number of active workers divided by the total steps. It amounts to 1.781 which is earned from dividing 57 (i.e. the sum of 40, 1, 1, 1, 1, 2 and 1) by 32 steps; and 5) the average durability of the protest is calculated from the sum of peak widths over the total steps. It is 0.313(=10/32).

Results

6.1 Figure 4 shows the up-and-down of the number of workers who participated in the protest without getting arrested at each step. The protest wave is ragged, sporadic, and rhythmic rather than smooth and regular, which is very similar to what we observed in the history (e.g. see Figure 1 and 2 in Oliver and Myers (2002)). Interestingly, a large size of uprising almost always occurs at the beginning. Its size is larger at WD=4 than the size at WD=2, and this trend becomes stronger, as less and less workers rely on tag-mediated perception. Taking a look at the slope of the surface plane, the overall number of active sites has the tendency to decline as the level of tag-mediated perception becomes low from the right to the left. The summary of descriptive statistics of five outcome variables is presented in Table 2.

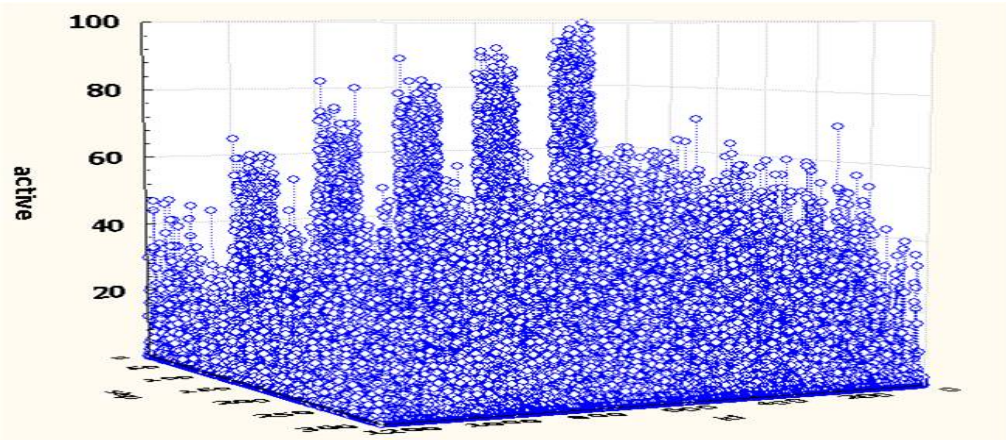

Figure 4. Change in the Number of Active Sites over Time

Notes: All cases are displayed along the experiment identification number in Table 1 (bottom front) is 1 through 12 from right to left. Another axis (bottom left) represents steps from 0 through 300.

Table 2. Descriptive Statistics of Outcome Variables

	N	Min	Max	Mean	Std.dev	Skewness	Kurtosis
Peak frequency	1200	5	79	52.65	14.563	-1.100	0.946
Waiting time	61980	0	125	2.570	4.428	5.842	69.649
Maximum peak's height	1200	2	103	38.51	27.157	0.553	-1.095
Average strength	1200	0.033	7.412	1.872	1.556	0.991	0.170
Average durability	1200	0.02	0.67	0.3681	0.169	-0.050	-1.240

Protest Cycle Frequency

6.2.1 There are two approaches to measuring the goodness of fit between the observed distribution and the theoretical distribution: the Chi-square test and the Kolmogorov-Smirnov test (KS test). The KS test has some advantages, but its range of applicability is more limited and its test significance is problematic especially when all the parameters of the hypothesized distribution are not known (Law, 2007). The occurrence of protests follow a Poisson distribution at WD=2 except one setting (Experimental ID=9) according to the Chi-square test[1] of the equi-dispersion requirement in the Poisson distribution – that the conditional mean of the outcome should be equal to its conditional variance. The frequency of protests never satisfies a Poisson distribution at WD=4. In other words, events are not independent at the higher level of wage dispersion (i.e. the higher correlation between the previous event and the next one).

Table 3. Chi-square Test of Fitting Peak Frequency to Poisson Distribution

Experiment Condition	1	2	3	4	5	6	7	8	9	10	11	12
Mean	48.83	61.12	48.72	62.28	48.93	63.40	48.22	66.09	44.64	66.31	16.60	56.66
Variance	41.19	26.61	53.15	21.88	55.88	23.52	59.04	28.95	50.56	30.38	16.81	48.81
Variance/Mean	0.84	0.44	1.09	0.35	1.14	0.37	1.22	0.44	1.13	0.46	1.01	0.86
df	16	13	18	13	19	23	19	12	23	14	21	16
P-value	0.258	0.001	0.551	0.001	0.109	0.008	0.072	0.001	0.000	0.002	0.547	0.022

6.2.2 As is shown in Figure 5, the overall trends are that: the more wage dispersion, the more frequent the protest; the more tag-mediated perception, the less peak frequency; and the interaction effects between wage dispersion and tag-mediated risk aversion. To test these effects, the negative binomial regression model[5] is recommended instead of the Poisson regression or the zero-inflated regression model because the average of peak frequency is much less than its variance and the number of 0s is small among the total 1,200 cases.

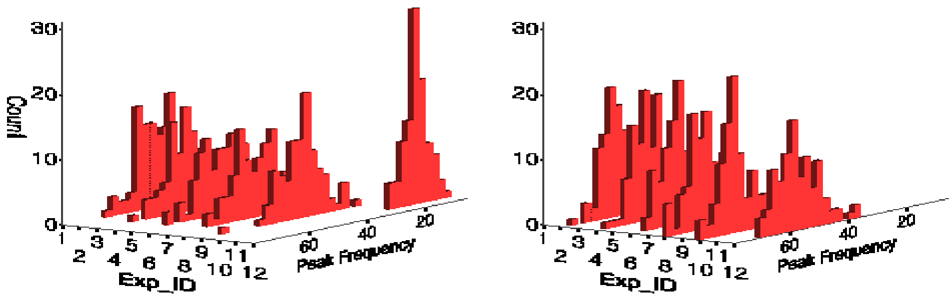

Figure 5. Histogram of Peak Frequency

Notes: Wage Dispersion=2 (Left); Wage Dispersion=4 (Right)

---

[5] Count variables are often treated as if they are continuous and the linear regression model is applied, but this makes estimates inefficient, inconsistent, or biased. The Poisson regression model is the most basic model among alternatives. However, "either unobserved heterogeneity or positive contagion can generate over-dispersion (Hannan and Carroll, 1992: 79)." The negative binomial regression model allows the variance to exceed the mean by incorporating an error term (i.e. unobserved heterogeneity) in the Poisson regression model (Long, 1997) although it cannot handle another issue that the occurrence of an event affects the rate of subsequent occurrences. Hannan and Carroll (1992) suggests quasi-likelihood estimation be an alternative approach to solve both autocorrelation and over-dispersion.

6.2.3 We do not require the linear relationship between tag-mediated perception and the outcome variable to prefer its dummy variables. Four dummy variables such as Tagdum1, Tagdum2, Tagdum2, and Tagdum4 are included given that 0% of tag-mediated perception is the reference category, besides the dummy variable of wage dispersion (i.e. Wdum, when the omitted category is WD=2). In Table 4, two models are specified – Model 1 without interaction terms and Model 2 with them. Provided with the difference between the log likelihood of a baseline model with the intercept only and the log likelihood of its full models, the tests of the global goodness-of-fit indicate that at least one of coefficients is not 0 in Model 1 and Model 2 ($p<.000$). As is similar in a family of logistic regression models, the incidence density ratio (IDR) helps better understand the substantive meanings of coefficients because it corresponds to the odds ratio. Also, $(e^b-1)*100$ expresses the percent change in the expected peak frequency with each one unit increase of explanatory variable.

6.2.4 In both models, 100% of tag-based risk aversion significantly decreases the average peak frequency by 35.3% ($p<.01$). However, the average peak frequency does not decrease either proportionally or sensitively to the increase in tag-mediated perception. On the other hand, the average peak frequency is about 50% higher at WD=4 than that at WD=2. The interaction effect between 100% tag-based risk aversion and the higher level of wage dispersion remains significant in Model 2 ($p<.01$). The effect size of 100% tag-based rational choice depends on the degree of wage dispersion: b is -0.100451 at WD=4, but it is -1.047182 at WD=2. In other words, the likelihood that underpaid workers participate in on-going local uprisings decreases drastically when they consider fully how many workers look trustworthy in their local areas given wage disparity smaller.

Table 4. Negative Binomial Regression of Peak Frequency

	Model 1			Model 2		
Variable	b	IDR	SE	b	IDR	SE
Tagdum1 (20%)	.0086931	1.008731	.0175757	.0095046	1.00955	.0134551
Tagdum2 (40%)	.0200983	1.020302	.0175453	.0214154	1.021646	.0134154
Tagdum3 (60%)	.0353365*	1.035968	.0175031	.0191893	1.019375	.0181148
Tagdum4 (80%)	.0022521	1.002255	.0175852	.0090546	1.009096	.0134566
Tagdum5 (100%)	-.4347205**	.6474456	.0190361	-1.047182**	.3509253	.0268919
Wdum (4)	.3986059**	1.489746	.0104814	.2809227**	1.324351	.0095829
Tagdum3*Wdum				.0343209	1.034917	.0212255
Tagdum5*Wdum				.9467431**	2.577302	.0295082
Intercept	3.794761			3.856585		
Pseudo-R2	0.1217			0.2254		
LR Chi2	1239.21**			2295.26**		
Df	6			7		

Notes: * $p < .05$ ** $p < .01$ (two-tailed tests)

Waiting Time between Protest Cycles

6.3.1 We checked whether the distribution of waiting time follows an exponential curve, which is parallel to the test of whether peak frequency meets a Poisson distribution[6]. In the context of survival analysis or event history analysis, the hazard function should be constant in an exponential distribution.

Table 5. Chi-square Test of Fitting Waiting Time to Exponential Distribution

ID	1	2	3	4	5	6	7	8	9	10	11	12
Mean	3.390	1.112	3.471	1.101	3.573	1.123	3.725	1.108	4.312	1.258	14.818	2.568
Stddev	3.800	1.534	3.970	1.552	4.030	1.535	4.352	1.516	4.837	1.738	14.803	3.050
Df	6	3	4	3	3	4	3	3	4	4	4	4
P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.086	0.000
N	4783	6012	4772	6128	4793	6240	4722	6509	4364	6531	1560	5566

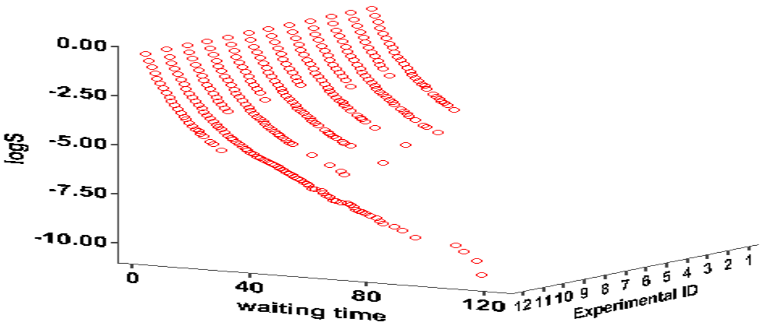

Figure 6. Log of Survival Function vs. Waiting Time

6.3.2 According to the Chi-square test, only one distribution of waiting time (Experimental ID=11) satisfies an exponential distribution (p=0.086). Putting together both the results in Table 3 and Table 5, the Poisson distribution and the Exponential distribution can be the best fit models of peak frequency and waiting time between neighboring peaks, respectively, only when workers take tag-based rational choice at the lower level of wage dispersion. This conclusion is also supported by the fact that the relationship between the log of Survival function, lnS(t), and waiting time is not linear except Experimental ID=11 (Figure 6). Recall that the relationship between the two is linear as long as the hazard function, h(t), is constant in the formula, $\ln S(t)=\int h(t)dt$.

[6] Students of probability models prefer the term – inter-arrival time to waiting time. Waiting time should be an exponential distribution as far as events happen rarely and independently (Moss, 1993).

6.3.3 As the next step of analysis, it is necessary to fit the observed waiting time into the Weibull model since the exponential distribution is one case of the Weibull model (i.e. Exponential(β) equals with Weibull(1,β). β is the shape parameter). One of properties in the Weibull distribution is well known as the linear relationship between the log of the negative log of Kaplan-Meier survival function and the log of time. Figure 7 demonstrates that the observed waiting time follows a Weibull distribution[7]. It should be noted that the distribution has negative time dependence[8], which implies that the chance of protest diminishes with time[9].

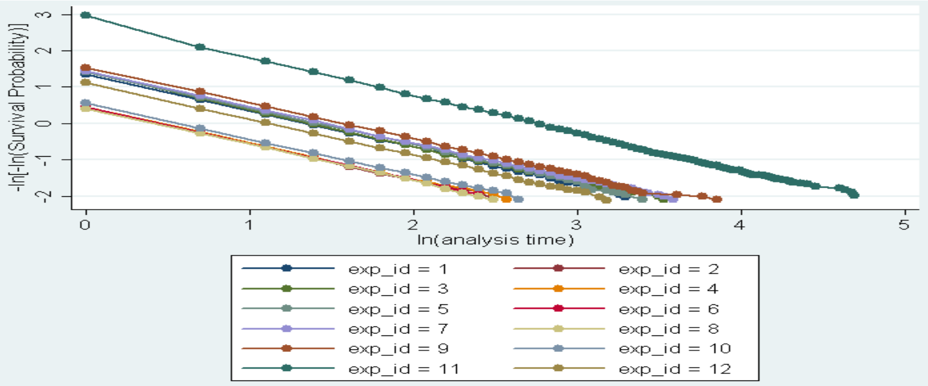

Figure 7. Fitting Waiting Time to Weibull Distribution

6.4 We employed bootstrapped median regression analysis hereafter – some other times called quantile regression or least absolute value regression) instead of robust regression analysis of the

---

[7] Suffice to say here that curves are parallel with each other to support the assumption of proportional hazard. It is not shown here, but the null hypothesis can be rejected that all survival curves are the same across the strata, according to both the log-rank test and the Wilcoxon test ($p<.01$).

[8] Epstein (2002) checked the shape of waiting time distribution for outbursts when the number of active sites exceeds 50. Without inferential statistical tests, he tentatively suggested a Weibull or perhaps Lognormal distribution (Ibid: 7245). It is not shown here, but the Lognormal distribution does not fit the observed waiting time. This is not surprising because the hazard function should increase and then decrease in the Lognormal distribution. He also fitted the truncated distribution of waiting time – after throwing away the cases with waiting time less than 30 cycles – into the negative exponential (R^2=0.98): "the negative exponential distribution is ubiquitous in the analysis of failure rates – the rates at which electrical and mechanical systems break down. It would be interesting if 'social breakdowns' followed a similar distribution (Ibid: 7246)." However, it does not seem that social breakdown has the memoryless property of an exponential distribution (Moss, 1993) that $Prob(T>a+b|T>a)=Prob(T>b)$ given two time points, a and b (i.e. the failure rate is not time-varying).

[9] Epstein (2002: 7247) addressed this issue: "Rather than chip away at the regime's legitimacy over a long period with daily exposes of petty corruption, it is far more effective to be silent for long periods and accumulate one punchy expose… they (incremental legitimacy reductions) incite the most aggrieved agents to go active prematurely, allowing them to be arrested by before they can catalyze the wider rebellion." This can be tested by incorporating the maximum peak's height into Model 1 in Table 4. It turned out that the IDR of the maximum peak's height is 0.9923 ($p<0.01$) – it reduces the frequency of protest only by 0.8%: the reason that the previous outburst suppresses the next one may not be due to a massive increase in layoff. We also observed that negative time dependence still exists in spite of almost the same distribution of grievance after a massive layoff.

conditional mean for the following reasons[10]: 1) the assumption of normality is violated for the maximum peak's height, the average strength, and the average durability, according to the Kolmogorov-Smirnov test; 2) the equal variance assumption is not met for these variables as stated by the Levene's test; and 3) there are no influential outliers on the report of Cook's d statistics. Data are re-sampled with 100 repetitions.

Maximum Protest Strength

6.5 The overall trends in Figure 8 are: the height of the maximum peak decreases as the level of tag-mediated perception increases on the one hand and as the degree of grievance due to wage dispersion becomes high on the other hand; and the maximum peak's height falls more significantly and sensitively when wage distribution is more unequal in terms of the interaction effects. The regression analysis in Table 6 substantiates these results. The effects of tag-mediated perception, wage dispersion, and the average perceived risk are statistically significant ($p<.01$). The interaction effects of tag-based rational choice with wage dispersion on the number of yet-to-be-caught protesters hold true except for its low levels (e.g. 20% and 40%).

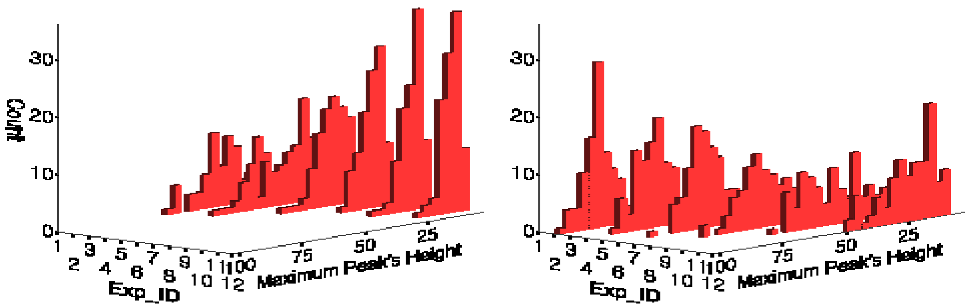

Figure 8: Histogram of Maximum Peak Height

Notes: Wage Dispersion=2 (Left); Wage Dispersion=4 (Right)

[10] It is not shown here, but analysis of covariance (ANCOVA) produced almost the similar results. This approach is, however, inappropriate given the non-normality of outcome variables although it can deal with covariates (i.e. the average risk) and interaction effects. One might point out that there are two nonparametric tests equivalent to one-way ANOVA: the Kruskal-Wallis test for between-subjects; and the Friedman's test for within-subjects (Field, 2006). It should be noticed that neither of them can handle two factors. In other words, the nonparametric technique equivalent to two-way ANOVA with interaction effects has yet to be developed (Sawilowsky, 1990). Since cases can be regarded independent, each case should contribute only one score (i.e. between-subjects). In the Kruskal-Wallis test, it can be assumed that a single factor has 12 levels in Table 1 (i.e. 6 from tag-mediated labeling and 2 from wage dispersion).We performed the KW test and the post-hoc Mann-Whitney test, but the results are pretty much the same with those from boostrapped median regression (e.g. Model 1 without either Average risk or interaction terms from Tagdum1*Wddum through Tagdum5*Wddum).

Table 6. Bootstrapped Median Regression of Maximum Peak's Height

	Model 1		Model 2	
Variable	b	SE	B	
Tagdum1 (20%)	-9.587855**	1.551989	-10..76169**	2.444639
Tagdum2 (40%)	-17.77437**	1.370096	-15.38132**	1.98836
Tagdum3 (60%)	-25.4557**	1.190489	-19.15335**	1.780269
Tagdum4 (80%)	-30.94983**	1.139763	-22.68818**	1.82074
Tagdum5 (100%)	-39.28033**	2.926962	-26.22719**	1.774393
Wddum (4)	41.3809**	.9360233	51.41315**	1.848912
Average Risk[11]	-231.4975**	42.12633	-123.2954**	24.91819
Tagdum1*Wddum			1.641277	2.690263
Tagdum2*Wddum			-0.9345663	2.158048
Tagdum3*Wddum			-10.23993**	2.358006
Tagdum4*Wddum			-21.64021**	2.762867
Tagdum5*Wddum			-38.24928**	3.288312
Intercept	156.3485		95.77074	
Pseudo-R2	0.6365		0.7243	
N	1200		1200	

Notes: * p < .05 ** p < .01 (two-tailed tests)

Average Protest Strength

6.6 The average height of peaks (Figure 9) exhibits a similar pattern as Figure 8: the more wage inequality, the more the average number of active workers; the more tag-mediated categorization, the smaller number of active ones; the more tag-based risk perception, the lower the average peak height; and the average number of active locals diminishes in proportion to the level of tag-mediated labeling and more quickly at WD=4 in spite of more rounded distributions. The bootstrapped median regression models support all of these results in Table 7. Besides the main effects (p<.01 for all), all interactions are statistically significant (p<.01) except the interaction of 20% tagging with WD=4 (p<.05).

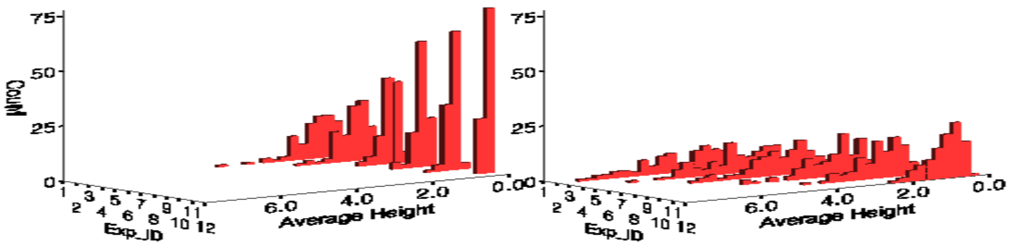

Figure 9. Histogram of Average Height
Notes: Wage Dispersion=2 (Left); Wage Dispersion=4 (Right)

[11] The average risk as a covariate in the models (Table 6 through 8) is the sum of rationally calculable Net Risk (N) across the population of active workers. In other words, there is no collinearity between the variable of the average risk and the dummy variables of tag-mediated perception. In contrast, we do not include the average grievance in the models afterwards because it is highly correlated to wage dispersion.

Table 7. Bootstrapped Median Regression of Average Height

	Model 1		Model 2	
Variable	b	SE	b	SE
Tagdum1 (20%)	-.3116859**	.0778818	-.2444202**	.0653775
Tagdum2 (40%)	-.6794687**	.0721597	-.4940914**	.057219
Tagdum3 (60%)	-.9856666**	.0693045	-.6190974**	.048365
Tagdum4 (80%)	-1.240048**	.0763999	-.8082274**	.0472705
Tagdum5 (100%)	-2.029495**	.1022127	-1.204908**	.0438649
Wddum (4)	2.063025**	.0365408	3.096444**	.177438
Average Risk	-29.97744**	2.637218	-17.93391**	1.933479
Tagdum1*Wddum			-.4373273*	.1910929
Tagdum2*Wddum			-.5308498**	.1964317
Tagdum3*Wddum			-1.184931**	.1936553
Tagdum4*Wddum			-1.416572**	.1886062
Tagdum5*Wddum			-2.186643**	.1775702
Intercept	16.66564		10.27003	
Pseudo-R2	0.6171		0.6882	
N	1200		1200	

Notes: * p < .05 ** p < .01 (two-tailed tests)

Average Protest Durability

6.7 The bootstrapped median regression analysis in Table 8 validates the overall patterns in Figure 10. First, the more tag-mediated risk aversion, the smaller peak width per step. In both models, the negative effect of tag-mediated perception on the average durability of protest is significant (p<.01) except when its level is 20% (p<.05). Second, the more wage disparity, the more durable workers protest. This positive effect is also significant (p < .01) in both models. Third, the interaction effects between the two are significant (p<.05 when 40% and 60% tag-based risk aversion comes into play; p<.01 when 80% and 100%). Interestingly enough, the chance that workers protest, once synchronized globally, persists decreases rapidly when workers are averse to taking risks based on the fullest consideration of how many workers look trustworthy in their local areas, regardless of the degree of wage dispersion between native and guest workers. Lastly, the effect of the average risk calculated rationally on the average durability of protest is not significant, which is in contrast with its significant effect on the average strength.

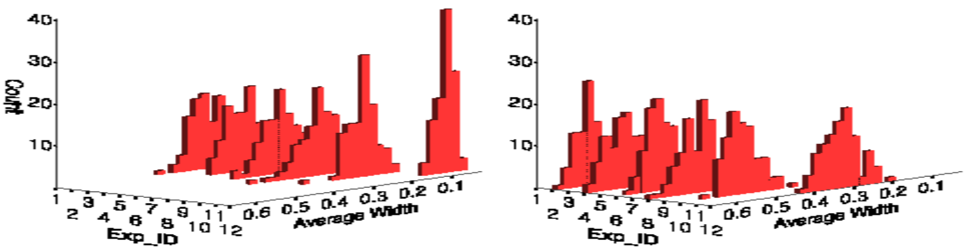

Figure 10. Histogram of Average Width
Notes: Wage Dispersion=2 (Left); Wage Dispersion=4 (Right)

Table 8. Bootstrapped Median Regression of Average Width

	Model 1		Model 2	
Variable	b	SE	b	SE
Tagdum1 (20%)	-.0110162*	.0053365	-.0174417*	.0076454
Tagdum2 (40%)	-.0241372**	.0065583	-.0378087**	.0071107
Tagdum3 (60%)	-.0475245**	.0058817	-.0567535**	.0074856
Tagdum4 (80%)	-.0806694**	.0054739	-.0891769**	.005871
Tagdum5 (100%)	-.2289275**	.0052425	-.2205993**	.0056
Wddum (4)	.2895023**	.0032918	.2827134**	.0078123
Average Risk	.0602099	.1233219	.1626939	.1372954
Tagdum1*Wddum			.0107118	.0110201
Tagdum2*Wddum			.0214555*	.0102841
Tagdum3*Wddum			.0126569*	.0097965
Tagdum4*Wddum			.0230848**	.0087773
Tagdum5*Wddum			-.0258576**	.0090211
Intercept	.2600106		.2132486	
Pseudo-R2	0.7815		0.7866	
N	1200		1200	

Notes: * p < .05 ** p < .01 (two-tailed tests)

Conclusion and Discussion

7.1 Whether protest events are independent is affected by both the degree of wage dispersion and the strength of tag-mediated group identity. The frequency of events and inter-occurrence time fit into a Poisson distribution and an exponential function when a sense of relative deprivation among workers is smaller and tag-mediated risk aversion works to the fullest. Otherwise, episodes of local uprisings are represented by recurrent events with negative time dependence in the form of the Weibull distribution of waiting time between adjacent cycles. How often the episode of protest emerges is not that sensitive to a wide range of tag-mediated similarity perception levels except underpaid workers take it into full consideration. Both how strong global protest is and how long it persists once synchronized are quite sensitive to mid-to-high levels of tag-mediated risk aversion, however.

7.2 Our agent-based model suggests that tag-based group identity plays key roles in collective actions. It would be partial to explain the dynamics of protest with a pure rational choice framework which focuses primarily on individuals – particularly with regard to its strength and durability. The current model can be advanced to explore the following theoretical questions: 1) roles of 'leaders.' They can be distinctive from the masses of workers in terms of vision, the maximum layoff term, or the reachability; 2) what if workers "learn diversity" over time, instead of bringing laid-off workers with the same tag back to the population?; and 3) what if “carrot” or wage flexibilization policies were introduced in addition to the “stick” policy of laying off protesters?

References

[1] BERK, RA (1974) A Gaming Approach to Crowd Behavior. *American Sociological Review,* 39(3), pp. 355-73.

[2] BONACICH, E (1972) A Theory of Ethnic Antagonism: The Split Labor Market. *American Sociological Review* 37, pp. 547-59.

[3] BRICHOUX, D and Johnson, P (2002) The Power of Commitment in Cooperative Social Action. *Journal of Artificial Societies and Social Simulation* 5(3) http://jasss.soc.surrey.ac.uk/5/3/1.html.

[4] EPSTEIN, JM (2002) Modeling Civil Violence: An Agent-based Computational Approach. *Proceedings of the National Academy of Sciences of the United States of America* 99(3), pp. 7243-50.

[5] FIELD, AP (2005) *Discovering Statistics Using SPSS*. 2nd Edition. London: Sage Publications.

[6] GRANOVETTER, M (1978) Threshold Models of Collective Behavior. *American Journal of Sociology* 83, pp.1420-43.

[7] HAMILTON, LC (2003) *Statistics with Stata: Updated for Version 8*. Belmont, CA: Thomson/BrooksCole.

[8] HANNAN, M and Carroll, G (1992) *Dynamics of Organizational Populations: Density, Legitimation, and Competition*. Oxford: Oxford University Press.

[9] HOLLAND, JH (1996) *Hidden Order: How Adaptation Builds Complexity.* Reading, MA: Addison Wesley Publishing Company, Inc.

[10] LAW, AM (2007) *Simulation Modeling and Analysis*. 4th Edition. NY: McGraw-Hill.

[11] LONG, JS (1997) *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks: Sage Publications.

[12] MARWELL, G and Oliver, PE (1993) *The Critical Mass in Collective Action: A Micro-Social Theory*. New York: Cambridge University Press.

[13] MOSS, SM (1993) *Introduction to Probability Models*. 3rd Edition. Boston: Academic Press.

[14] OLZAK, S (1989) Analysis of Events in the Study of Collective Action. *Annual Reviews in Sociology* 15, pp.119-41.

[15] OLIVER, PE and Myers, DJ (2002) The Coevolution of Social Movements. *Mobilization: An International Journal* 8(1), pp.1-24.

[16] RIOLO, R, Cohen, MD, and Axelrod, R (2001) Evolution of Cooperation without Reciprocity. *Nature* 414, pp. 441-3.

[17] ROHLINGER, DA and Snow, DA (2003) "Social Psychological Perspectives on Crowds and Social Movements." In Delamater J (Ed.) *Handbook of Social Psychology,* New York: Kluwer Academic/ Plenum Publishers.

[18] SAWILOWSKY, SS (1990) Nonparametric Tests of Interaction in Experimental Design. *Review of Educational Research* 60(1), pp. 91-126.

[19] TAJFEL, H (1974) Social Identity and Intergroup Behavior. *Social Science Information* 13, pp.65-93.

[20] TILLY, C (1978) *From Mobilization to Revolution*. MA: Addision-Wesley.

[21] WATTS, DJ (2003) *Six Degrees: The Science of a Connected Age*. New York: Norton.

[22] WILENSKY, U (1999) NetLogo. Evanston, IL: Center for Connected Learning and Computer-Based Modeling. Northwestern University. http://ccl.northwestern.edu/netlogo.