A Robust Measure to Uncover Community Brokerage in Illicit Networks

Masarah Paquet-Clouston

m.paquet-clouston@umontreal.ca

School of criminology, Université de Montréal

2900 Edouard Montpetit Blvd

Montreal, Quebec H3T 1J4

+1 514 601 9243

Martin Bouchard

mbouchard@sfu.ca

School of criminology, Simon Fraser University

8888 University Drive,

Burnaby, BC CANADA, V3E 1S6

Published in 2022 in *Journal of Quantitative Criminology*

DOI: 10.1007/s10940-022-09549-6

Acknowledgements. The authors want to thank the research analysts who collected the data, and the criminal intelligence agency for sharing them with us. Thank you to Serge-Olivier Paquette, Mitch Macdonald, the Editor, and reviewers of JQC for their comments and suggestions made on earlier versions of this manuscript.

A ROBUST MEASURE TO UNCOVER COMMUNITY BROKERAGE IN ILLICIT NETWORKS

ABSTRACT (250 words)

Objectives: Brokers are said to be the oiling chain of illicit networks, facilitating the efficient flow of illicit products to destination. Yet, most of the available brokerage measures focus on local or individual networks, missing the brokers who connect others across communities, such as market levels. This study introduces a robust measure that uncovers, scores, and positions these community brokers.

Methods: We used network data aggregated from numerous investigations related to 1,800 criminal entrepreneurs operating in Western Canada. After uncovering the communities using the Leiden algorithm, we developed a community brokerage score that assesses individual potential reach and control at the meso level, and that accounts for individual position changes due to different community structures. We examined how the score relates to brokerage and structural hole measures as well as seriousness of involvement in criminality.

Results: We found that the illicit network studied has a strong and stable community structure, and community brokers form about 9% of the population. The score developed is statistically robust and is not strongly related to network and structural hole measures, which confirms the need for a novel measure that captures this strategic position in illicit and other networks.

Conclusions: Community brokers are especially important in illicit networks where large-scale covert coordination among criminal entrepreneurs is risky. The measure we propose is not overlapping with currently existing brokerage measures and has the potential to contribute to our understanding of how products and information flow beyond local networks, in criminology and other fields.

Keywords: Illicit networks, community detection, illegal drug markets, brokerage,

1. INTRODUCTION

A relatively robust finding of network research is that individuals who bridge relationships between otherwise unconnected others – brokers – tend to fare better, be it through controlling information flows (Freeman, 1977), obtaining better opportunities (Granovetter, 1973 1983) or simply having higher social capital (Burt, 2004; 2009). This finding has been observed in illicit networks as well, with criminal entrepreneurs located in brokerage positions actively engaged in leveraging their comparative advantage in having longer criminal careers (Morselli, 2001; 2003), making more money from crime (Morselli, Paquet-Clouston and Provost, 2017; Morselli, 2009a, b), or avoiding arrest (Morselli 2010).

Beyond the individual advantages that brokers may get from their strategic position, a striking feature of brokerage is the role it appears to play at the meso level, in connecting individuals belonging to different criminal groups (Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017; Schaefer et al., 2017), or market levels (Bright, Koskinen and Malm, 2018; Malm and Bichler, 2011). Illicit drug distribution chains are relatively long, making it almost impossible (and especially risky) for a single group to connect all steps of the trade. Instead, criminal entrepreneurs forego some efficiency for the security of additional layers of distribution, relying on brokerage to connect the chain (Bright, Koskinen and Malm, 2018). Brokers are thus essential in making illicit deals happen, using their connections to bridge otherwise loosely knit networks of criminal entrepreneurs (Morselli and Roy, 2007; Morselli, 2001a, b; 2003; Reuter and Haaga, 1989). Their position makes them high-value targets from a law enforcement perspective and removing them has been found to effectively disrupt illicit networks (Duxbury and Haynie, 2019; Bright et al., 2017; Bright, 2015; Malm and Bichler, 2011; Morselli and Roy, 2008).

Extant research in illicit networks is focused on the benefits of brokerage at the individual level, which is often calculated with Freeman's betweenness centrality measure (Morselli, Paquet-Clouston and Provost 2017; Morselli, 2009a, b). However, brokers can also be conceptualized at the meso-level, as individuals positioned *between* communities. This is especially relevant considering that numerous ethnographies of high-level dealers report the potential meso-level power of community brokers (Zaitch, 2002; Adler, 1993; Klerks, 2001; Reuter and Haaga, 1989; Desroches, 2007). Moreover, to uncover higher-order structures within social networks, various community detection algorithms have been developed (Dao, Bothorel and Lenca, 2020; Lancichinetti and Fortunato, 2009) and these algorithms can be leveraged to detect such meso-level brokers. However, when studies do examine larger social structures using various community detection algorithms, they do not measure meso-level brokerage power.

We believe this lack of meso-level research on brokerage in illicit networks is due to two reasons. First, illicit network scholars generally lack data beyond a single criminal investigation or case study that would be amenable to uncover brokerage beyond a specific group. Several studies investigated brokerage on its own, without examining the specific brokers across communities (e.g., Morselli, Paquet-Clouston and Provost, 2017; Morselli, 2009a, b). The reverse situation – meso-level studies of communities within illicit networks – are also common (Ouellet, Bouchard and Charette, 2019; Lantz and Hutchison, 2015; Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017; Schaefer et al., 2017). Yet, most of these studies do not focus on the individuals who may contribute to linking communities together. Both Ouellet, Bouchard and Charette (2019) and Lantz and Hutchison (2015), for example, made important meso-level contributions by examining how networks of multiple criminal entrepreneurs overlapped to form communities that were not necessarily following the boundaries expected from non-network data sources. Yet, the objective of these studies was not to identify the brokers that held these communities together. The few studies that examined meso-level brokers in illicit networks focused on describing some of their personal attributes (e.g., rank in an organization) compared to others (Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017; Schaefer et al., 2017). These studies, however, did not rely on measures that could detect and score meso-level brokers systematically.

Second, even when the data are available, the existing measures are not necessarily suitable to the task. Brokerage measures, in the tradition of Burt's (2009) structural hole theory, are meant to describe individuals in their immediate environment, not how these individuals fare as the glue that holds the larger social structure together. Another classic measure, Freeman betweenness centrality (1977), does better at uncovering brokerage within one large network, but it cannot distinguish brokerage at the meso level across large central communities. This is because the measure counts the number of times an individual sits on the shortest path between all pairs of individuals in the network (Freeman, 1977) regardless of the communities in which the individual belongs. This gives more weight to individuals connecting further clusters than those central to the network. To consider meso-level brokerage, community structures have to be detected prior to calculating brokerage. Alternatively, Jones, Ma and McNally (2019) developed two interesting meso-level brokerage measures to study a comorbidity network: bridge strength and bridge betweenness. While these measures come close to filling the gap in identifying and scoring community brokerage, none fully matches our needs. The first one only calculates the number of times an individual sits between two communities, and the second one is almost a replication of the betweenness centrality measure. There is a need to develop one global measure that assesses brokers' control over products and information flows simultaneously across communities.

In this study, we use a network of 1,800 criminal entrepreneurs to develop a robust measure that quantifies meso-level brokerage, which we name community brokerage. We use the term 'entrepreneur' to broadly capture the fact that all of these individuals are involved in continued illegal drug business for profit[1]. As an initial step, we uncover the illicit network's various community structures provided by the Leiden algorithm, which focuses on finding dense clusters in a network (Traag, Waltman, and van Eck, 2019). Rather than choosing one community structure uncovered by the algorithm, we consider thousands of the algorithm's iterations to find all possible community brokers. This procedure allows us to find that community brokers, in this network, form 9% of the population in all community structures (also named "network partitions") uncovered.

Then, a measure of community brokerage is developed and presented. For each partition found through the Leiden community detection algorithm, a local community broker score is calculated. This local score quantifies, for each bridge created, the bridge's size (the number of people connected through the bridge), efficiency (how far and how easily these people can be reached (i.e., cohesion) and exclusivity (whether other brokers connect these two communities). The results are summed up to calculate community brokers' reach and control at the meso-level for one partition. By *reach*, we mean the extent to which the individuals in the communities bridged can easily be accessed by brokers; by *control* the extent to which a community broker has exclusivity over a bridge – how many others play the same structural role in the community? The global community broker score is an average of all local scores, making it *robust* to the

[1] In this paper we purposefully avoid the term "organized criminals" because its variety of definitions may detract readers from the main profit-oriented activities at the heart of their inclusion in the sample.

inherent randomness of community partitioning. The averaged global score thus follows the partition distribution found when running the community detection algorithm thousands of times. This implies that a partition that emerges more often has more weight than an outlier partition (although the outlier partitions are still considered). How the community broker score differs from other brokerage measures (including Freeman betweenness centrality, bridging centrality, structural hole measures, bridge strength and bridge betweenness) is discussed throughout the results, using non-parametric tests and pedagogical examples. We conclude that the differences found illustrate that the score taps into a different brokerage pattern: *community brokers reach and control at the meso level*.

The discussion section illustrates the flexibility and range of the measure developed: the *focus* can be changed for access to various strategic products (as opposed to the number of people), such as the search for firearms. It can also be calculated based on other community detection algorithms than the Leiden algorithm, like the Girvan-Newman algorithm, which focuses on partitioning networks based on edge betweenness. A Python open-source library is available for researchers who wish to further explore meso-level brokerage[2]. This study pushes the brokerage discourse beyond the individual level to include higher-level structural forms. It opens new research avenues while providing tools needed to conduct such meso-level inquiries.

2. MEASURING BROKERAGE AT THE MESO LEVEL

2.1 BROKERS AND ILLICIT NETWORKS

At the onset of network science, brokerage positioning quickly garnered the attention of scholars (Bavelas, 1948; Shimbel, 1953). Those positioned as "brokers" are defined as individuals who bridge "gaps in communication between persons, groups, structures and even

[2] Information about the package can be found here: https://github.com/Masarah/community_broker_score

cultures” (Boissevain, 1974, 148). By bridging gaps or connecting clusters, brokers act as points of controls for information flows in the social network they evolve (Freeman, 1977). Such strategic positions give them information diffusion power: they can accelerate, hinder or change (bias) the information that circulates in a network (Burt, 2004; 2009). In illicit networks, brokers also control not only information flows, but also product flows: they often connect various groups of the supply chain (Bright, Koskinen and Malm, 2018; Malm and Bichler, 2011).

Existing research investigating brokers in illicit networks can be broken down into studies examining the individual broker, including the personal costs and benefits of holding this sort of position, and studies examining brokerage as a position of interest in a larger social structure. In terms of personal costs and benefits, brokers were shown to typically do better, whether measured in criminal earnings (Morselli and Tremblay, 2004; Morselli, Provost and Paquet-Clouston, 2017), or their ability to avoid arrest (Morselli, 2010). Strategically building one’s personal network towards a greater level of non-redundancy (i.e., positioning yourself as a broker) was also shown to increase status and extend the careers of criminal entrepreneurs (Morselli, 2001a, b; 2003). In terms of brokers as key positions of interest in larger networks, brokers were found to represent a network vulnerability as they are more exposed and thus face higher risks of being arrested (Malm et al., 2017; Calderoni, 2011). From a law enforcement perspective, brokers are considered to be a key position for network disruption. (Duxbury and Haynie, 2019; Bright et al., 2017; Bright, 2015; Malm and Bichler, 2011; Morselli and Roy, 2008). However, such effect may be short-lived as the remaining members of the network often quicky adapt and replace the removed individuals (Bouchard, 2007; Bright, 2015; Duijn et al., 2014; Morselli and Petit, 2007).

Yet, the current study is not focused on what illicit networks do for brokers (e.g., higher criminal earnings) or how brokers can be leveraged to disrupt them; rather, we are interested in *what brokers do for networks*. Brokers are considered to be the conduits through which otherwise unconnected groups can collaborate. In the context of illegality, larger social structures become challenging to create, and maintain (Reuter, 1983). Criminal groups are generally small, but the challenge of illicit product distribution remains. A certain form of cooperation between different groups is needed to overcome these challenges and brokers fill that role (Bouchard and Morselli, 2014).

In the context of illicit networks, this is often referred to as the trade-off between structuring networks to favor efficiency (i.e., open and direct connections across a relatively short distribution chain) and security (i.e., multiple relatively independent cells that are loosely connected in a relatively long distribution chain) (Morselli, Giguère and Petit, 2007). Although there are contradictory conceptualizations of the efficiency and security tradeoff within the illicit network literature (Bright, Koskinen, and Malm, 2018), the trade-off is often solved with brokerage.

Brokers have been found to control the flow of information or products between groups, while maintaining network's flexibility and playing a valuable role in connecting these groups together (Bright, Koskinen, and Malm, 2018; Morselli, 2009b; Morselli and Roy, 2008). For example, Bright, Koskinen, and Malm (2018) investigated an Australian drug trafficking network and found that individuals preferred to protect themselves by using "brokers that offer access to the rest of the network" (p.237). Actors would maximize their efficiency and maintain their secrecy by being tied to actors within their own social networks (their local clusters) and minimize the number of external ties to the cluster. Brokers are those connecting other disconnected clusters, and this implies that those who use brokers do not have to take these additional risks themselves.

In sum, brokers are the oiling chain of illicit networks: they allow groups of individuals to be connected, providing network flexibility and efficiency. Although brokers are key in illicit networks, their power depends on the structure of these networks (Bright and Whelan, 2020). What does the larger social structure of communities that need brokers look like? And what is the role of a broker positioned in-between communities?

2.2 COMMUNITY STRUCTURES AND MESO-LEVEL BROKERS IN ILLICIT NETWORKS

The structure of social networks, licit or illicit, may be assessed at the meso level by first testing whether networks can be broken down into clusters of individuals that are connected. These groups of highly connected nodes are known as *communities* (Fortunato, 2010; Hu et al., 2008; Newman, 2006), but also clusters, cohesive groups, or modules (Palla et al., 2005). There has been an increasing body of literature studying the mesoscopic aspect of networks and various community detection algorithms have been developed to automatically detect community structures in networks (Dao, Bothorel and Lenca, 2020; Lancichinetti and Fortunato, 2009).

For example, the fast greedy (Newman, 2004b), the Louvain (Clauset et al., 2004) and the Leiden (Traag et al., 2019) algorithms find communities via the optimization of a cost function (i.e., modularity) and aggregation of communities iteratively. The Girvan-Newman algorithm partitions a network via an edge betweenness measure. The InfoMap algorithm, on the other hand, aims at maximizing compactness while minimizing information losses when finding communities in a network (Rosvall and Bergstrom, 2008). The Markov cluster algorithm simulates a diffusion process with expansion and inflation parameters until an equilibrium in communities detected is found (Dongen, 2000). The Walktrap algorithm simulates a random walker to find communities, assuming that random walkers are likely to stay within a cohesive set of nodes (Pons and Latapy, 2006). Such algorithms have been used to understand community structures and group dynamics in various illicit network contexts as well as to flag those with strategic positioning at the mesoscopic level.

Recently, two studies on illicit networks used network-based communities as a substitute for traditional groupings of offenders and examined whether communities of different sizes and density were associated with an outcome of interest (Ouellet, Bouchard and Charette, 2019; Lantz and Hutchison, 2015). For example, Ouellet, Bouchard and Charette (2019) used the Louvain algorithm to detect criminal groups in Montreal, testing whether specific network characteristics were associated with survival. They found that group survival is a function of group cohesion and embeddedness. Lantz and Hutchison (2015), on the other hand, used the Girvan-Newman algorithm to detect groups within a network of sentenced burglary offenders and assessed the characteristics and the duration of groups' offending careers, the impact of group membership on individual criminal careers and the impact of arrest on the offending patterns of connected co-offenders. They found that larger groups with dispersed offending structure are more likely to survive.

Other studies focused not so much on communities themselves, but like us, on the individuals that help connect them together. For example, DellaPosta (2017) used community detection on an American mafia network, showing that bridging connections are concentrated among a small group of individuals who tend to either be of high, or low status in their respective mafioso family (instead of a "middling" status). The author hypothesized that those with lower status in a criminal organization may be more inclined to have boundary-spanning bridging position because they have nothing to lose, while those with higher status maintain that position because there are insulated from peer-judgments. Individuals in the middle are more likely to “conform” to their immediate peers, and thus avoid contact with other mafioso families.

The Louvain method was used by Calderoni, Brunetto and Piccardi (2017) to compute a community analysis of the ‘Ndrangheta. The authors found that the subgroups uncovered in the analysis (based on social interactions among members) generally reproduced the expected internal organization of the ‘Ndrangheta . They also measured intra-community (an individual’s level of connectivity in a community) vs inter-community (an individual’s level of connectivity to other communities) strength for each individual, revealing mafia “bosses” tended to have higher intra- and inter- connectivity scores. Important for our purposes, these results imply that uncovering key actors *within and across communities* is a potentially powerful proxy to determine criminal leadership – a piece of information that may not otherwise be known to researchers. These results are not incompatible with those of Schaefer et al. (2017) who applied the Louvain algorithm to a prison network. They found that 19 out of 131 individuals (14.5%) connected various communities together and these meso-level brokers tended to have more ethnically and racially diverse networks than others.

2.3 MEASURING MESO-LEVEL BROKERAGE POWER

As shown in previous studies on illicit networks, looking at higher order structures unveils further information on the organization of a network (Gulbahce and Lehmann, 2008; Palla et al., 2005). From such information, questions arise as to how brokers involved in bridging these clusters are detected with standard measures and whether studying brokerage power at the mesoscopic level may reveal previously unknown information on key individuals. None of the previously cited studies relied on such as measure.

Traditionally, brokerage power has been quantified with *Freeman betweenness centrality*, a measure that counts the number of times an individual sits on the shortest path between all pairs of individuals in the network (Freeman, 1977). Hwang et al. (2006) also developed a *bridging centrality* measure to quantify the potential power of a node to control flows of information in a network. Bridging centrality is calculated by multiplying a node's bridging coefficient (which measures the extent to which a node is connected with nodes of high degree) with a node's betweenness centrality score. However, in both cases, the measures cannot distinguish brokerage at the meso level across large central communities. This is because betweenness centrality (part of the bridging centrality score) counts the number of times an individual sits on the shortest path between all pairs of individuals in the network (Freeman, 1977) and thus give more weight to individuals connecting further clusters than those central to the network.

Alternatively, Burt (2009) developed the structural hole concept to capture those who sit on network gaps and can *broker* otherwise unconnected segments, thus controlling information diffusion in the network. To capture this brokerage power, three measures related to an individual's direct network (ego network) were developed (Burt 2009): *effective size, efficiency,* and *constraint*. Effective size is based on the concept of redundancy: if a person's contacts are connected to one another, then the contacts are redundant. Efficiency, on the other hand, represents the effective size divided by the actual size of the person's network and shows how efficient a person's network is regardless of its size. Constraint is an indicator of the extent to which a person's connections are connected to other individuals in a network. Higher constraint means that the person is limited in terms of social interaction reach. Although interesting, Burt's measures focus on an individual and their direct network: whether an individual's network is efficient in terms of having non-redundant ties. It does not consider the potential of a network as a whole.

Past studies that have investigated illicit networks at the meso-level have found that they follow community structures (Ouellet, Bouchard and Charette, 2019; Lantz and Hutchison, 2015; Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017). Hence, studying those who bridge communities can yield valuable insights on key actors positioned at the edge of communities. Granovetter's (1973) seminal work on the strength of weak ties stresses the power of weak ties *in-between groups*. An individual with many weak ties, known as a "bridge" or a "liaison person", is best placed to diffuse information due to the network position. Weak ties refer to ties outside the immediate circles of peers and extend to informal opportunities (Portes, 1998). Thus, if we conceptualize that opportunities emerge from ties *in-between* clusters of individuals, then we should try to find those individuals who bridge *clusters*, considering that such groups may be beyond their direct network. *Bridges between clusters* become a function of brokerage opportunities that can be grasped neither at the macro nor micro levels, but solely through a mesoscopic perspective.

Recently, Jones, Ma and McNally (2019)'s developed two measures that aimed at grasping this meso-level brokerage power for a comorbidity network, *bridge strength* and *bridge betweenness*. Bridge strength counts the number of times an individual sits between two communities and bridge betweenness assesses the number of times an individual sits on the shortest paths between two nodes from different communities. Both measures were found to be efficient at finding key nodes in the network at the meso level.

Yet, these measures do not focus on quantifying community brokerage, but rather finding strategic nodes. For example, bridge strength provides equal strength to brokers for each community bridged, regardless of the *size* and cohesion of the communities. Bridge betweenness considers bridge *size* and *exclusivity*, by favoring one broker over the other depending on who sits at the shortest path between two communities. Yet it does not penalize equally two individuals who are brokering the same communities, thus diminishing their control over the bridge. Moreover, calculating these measures require that the community structure of the network be known in advance, or a specific network partition be (sometimes arbitrarily) chosen, as community detection algorithms developed in network science may yield various network partitions that are equally good.

Previous studies have showed the importance of brokers in illicit networks (Morselli, 2009a, b) and the presence of community structures in these networks (Ouellet, Bouchard and Charette, 2019; Lantz and Hutchison, 2015; DellaPosta, 2017; Calderoni, Brunetto and Piccardi, 2017; Schaefer et al., 2017). Although there exist brokerage measures, none focused on quantifying community brokerage: the extent to which an individual has control over information and product flows at the meso-level. Moreover, with more and more community detection algorithms being developed to automatically uncover community structures (Dao, Bothorel and Lenca, 2020; Lancichinetti and Fortunato, 2009), as well as the theoretical importance of individuals connecting communities, the next logical step is to quantify community brokerage.

Leveraging a network of 1,800 criminal entrepreneurs, the study's objective is to develop a robust, flexible, and all-encompassing measure of community brokerage using community detection algorithms. The measure developed sums, for each bridge created by a community broker, the bridge's size (i.e., the number of people reached), efficiency (i.e., how easily individuals can be reached in the community (cohesion)) and exclusivity (i.e., how many other brokers are bridging these same communities). It is flexible by design and considers many network partitions, thus controlling for the potential variations in network structures derived from different community detection algorithms. The measure is compared with other brokerage measures via a series of non-parametric tests and case studies.

3. DATA AND METHODS

3.1 DATASET

This study uses a dataset depicting the social network of individuals involved in drug trafficking and drug market distribution in a Western Canadian Province. The data used in this study was collected by a team of analysts from the province-level criminal intelligence agency, as part of the service's annual Provincial Threat Assessment reports. The intelligence agency acts as a central hub for strategic analysis and intelligence sharing on organized and serious crime in the province and focuses on collecting information on serious and organized crime from the various police agencies activein the province.

3.1.1 Dataset Creation

The team of analysts, in partnership with an academic network researcher who is also a co-author of the current manuscript, developed a coding procedure for the sole purpose of conducting social network analyses. The social network, which was completed in 2019, depicted relationships (one-mode network) among individuals who were part of a police investigation related to illegal drug market supply over the course of two years: 2017-2018. During these two years, investigators in charge of a case involving such criminal entrepreneurs[3] were asked to a submit report to the intelligence analysts with a summary of the case and of the connections among the persons of interest in the investigation, including their contexts. The analysts then

[3] We use the term 'entrepreneur' to broadly capture the fact that all of these individuals are involved in continued illegal drug business for profit.

took the reports and searched police databases for any other cases involving these individuals *within the 2017-2018 window period*. The aim was to uncover any and all types of known connections among individuals. These ties were coded into a standard edge list of individual to individual connections and their type (social, criminal, business, familial, unknown). Police databases included: (1) case files (official documents and legal evidence with relationship information), (2) street checks (police-citizen contact when individuals were seen together) (3) surveillance field notes (relationships observed during physical surveillance by officers), and (4) informants (relationships reported by police informants). The data thus contain more information on criminal associations than police data strictly based on co-arrests. As these data sources were searched for, new cases involving other individuals were coded and added to the social network, meaning that individuals can be connected to others beyond the original report that made them part of the dataset. Analysts aimed to only keep individuals suspected to be involved in illegal drug market activities. Their goal, in other words, was to build a database of criminal entrepreneurs for intelligence purposes[4]. All analysts who were involved with coding were trained on social network analysis by an academic network researcher who is also a co-author of the current manuscript. Analysts were trained for network coding on a similar set of investigative files to improve inter-coder reliability[5].

Each one-to-one relationship (i.e., tie) was coded as either social, criminal, familial, business, or unknown. Social ties represented connections with a social (or "non-criminal") purpose, such as two individuals seen in a bar, while criminal ties represented relationships where individuals were seen together in a criminal context. Family ties represented ties where

[4] We suspect that a small number of individuals included had a very minor role, or no role at all in the trade, especially among the individuals who only show a single connection in the network. None of these individuals can be identified as brokers.
[5] Unfortunately, a measure of inter-coder reliability was not calculated before the dataset was sent to the researchers.

two individuals came from the same family, such as a brother or a cousin, while business ties represented relationships that involved a business transaction. Over 89% of relationships were either coded as criminal or social and only a small number were coded as familial, business, or unknown relationships. All of these types of relationships were merged into a binary tie denoting the presence or absence of a tie. The important consideration, for our purposes, was to avoid creating missing or incomplete data that would create false positives – community brokers who, in fact, were not–, in situations where all types of relationships were not considered. The main criterion was that individuals knew each other, so all types of relationships were treated as a single network.

3.1.2 Sample

The raw social network we obtained had 2,868 individuals and 6,092 ties between them. First, we removed a few large cliques that were outliers: they were the results of coding interactions occurring from events where all individuals were observed together (e.g., parties, funerals). To keep our focus on the one-on-one interaction described in most of the data, which was the meaningful backbone of the network, we removed the ties found in five perfectly connected cliques in the network. No individuals were removed from this step, all of them had other connections in the network. Second, the network was structured as one large component and smaller networks disconnected from the main component. Since this study aims at identifying communities, we focused only on the largest connected component of the network. Nodes that are not connected to the main component would simply end up in stand-alone communities, with no community brokers. Focusing on the largest component reduced the network to 1,803 individuals and 3,527 edges.

3.1.3 Attributes Data

Table 1 lists the individual characteristics that were available to us in the data.

[INSERT TABLE 1 ABOUT HERE]

Beyond age and gender, the other six attributes are binary variables capturing various ways an individual may be involved in crime: 1) whether a person has an FPS number (meaning that the fingerprints of the individual are recorded in the national criminal database), 2) whether the person was previously known to be engaged in criminality, 3) whether the person is flagged armed and dangerous, 4) whether the person is flagged as violent, 5) whether the person is prohibited from having a weapon and 6) whether the person has previous weapon charges. We did not have a priori hypotheses as to whether community brokers would score high or low on any of these items separately[6]. Our interest was to determine the brokers' general level of involvement in crime and violence. To measure this, we summed the variables, creating an index, called *severity score*, ranging from zero to six, with six representing high involvement in criminality. The Cronbach's alpha coefficient was 0.75, indicating a good level of index reliability among the six variables.

3.2 EXTRACTING COMMUNITY STRUCTURES

As mentioned above, a plethora of community detection algorithms have been developed and all of them have their own specificities and optimization algorithms (Dao, Bothorel and Lenca, 2020; Lancichinetti and Fortunato, 2009). After reviewing the many common community detection algorithms used in social sciences (Leiden, Louvain, Girvan-Newman, Walktrap and

---

[6] This is especially important as some of the original scoring of individuals is dependent on investigators or analysts' coding practices, such as whether an individual is known to engage in crime or flagging someone as armed and dangerous. Combining the attributes allowed us to get a better sense of the scale of what is known about the seriousness of criminal involvement of each individual.

Infomap), we concluded that the most appropriate algorithm to quantify community brokerage for this study is the Leiden algorithm. It focuses on finding *non-overlapping dense clusters* by optimizing a modularity function through an iterative method (Traag, Waltman, and van Eck, 2019). Focusing on non-overlapping dense cluster segments the network in cohesive communities. The denser the communities, the more easily community brokers can exploit their control over information and product flows, by reaching individuals in the bridged communities quickly. The algorithm also does not require prior knowledge of the network's partition and necessitate no parameterization of the detection function.

Another algorithm that focuses on finding dense clusters is Louvain, the most widely used algorithm in criminology (Calderoni, Brunetto and Piccardi, 2017; Ouellet, Bouchard and Charrette, 2019; Schaefer et al., 2017). These two algorithms are related as *Leiden* is a recent improvement over *Louvain*. Traag et al. (2019) showed that Leiden corrects for Louvain's tendency to find internally disconnected communities while being faster at finding dense clusters. In practice, our preliminary results show that both algorithms largely overlap in their solutions for the social networks we analyzed.

One additional element that is worth noting is that the choice of algorithm largely depends on researchers' topic of study. We chose the Leiden algorithm because it was more suited to our theoretical framework of community brokers having control over information and product flows in illicit networks. However, other algorithms can be better suited, depending on researchers' aim when measuring community brokerage. The Python package that will be made available will allow such flexibility. Researchers will be able to test different algorithms prior to computing the community broker score, ultimately selecting the one that better fit their data and study aims.

Specifically, we used the **Leiden package**[7] with the *modularity vertex partition type* as the optimization cost function to uncover the community structures in the network. This cost function measures how good a division of a graph is in respect to another random division of the same network (Newman, 2006). It is "proportional to the difference between the number of edges within communities and the expected number of such edges" (Gulbahce and Lehmann, 2008, p.935). The expected number of edges, the base network, is a random graph with same degree and sequence (Fortunato and Barthelemy, 2006). The score is strictly less than one and increases if there are more edges between nodes than what would be expected in the random base network. This cost function, however, suffers a resolution limit: it is unable to detect small communities in large networks (Fortunato and Barthelemy, 2006). This study leverages a network of 1,800 entrepreneurs to find community brokers and quantify their brokerage power, detecting *smaller* communities is thus not as important as separating the network into large dense clusters, even if smaller communities may prevail within each of them.

Graph Partitions and Comparisons

Due to the modularity cost function, different instantiations of the algorithm returned different community structures. For this reason, we ran 50,000 trials on the network and made sure that the statistical distribution of the uncovered partitions was stable. We found a total of 487 unique partitions of 37, 38, 39 and 40 groups with different modularity scores. To compare the partitions, we computed the *Adjusted Rand Index* for each partition pair using the *AdjustedRandIndex* function in the *R mclust package*[8]. An index score zero means that two

[7] https://pypi.org/project/leidenalg/
[8] https://www.rdocumentation.org/packages/mclust/versions/5.4.7/topics/adjustedRandIndex

partitions are completely different, and a score of one means the two partitions are in perfect agreement. The results section presents how the different partitions differed based on this index.

3.3 FINDING AND SCORING COMMUNITY BROKERS

Once community structures were found, we focused on the criminal entrepreneurs who bridged them, the individuals we label as community brokers. A bridge is an edge between two individuals who belong to two different communities. A community broker is thus an individual who bridges at least two different communities. However, a binary variable “community broker/non-broker” is insufficient to understand a community broker’s value and position. For instance, it does not indicate whether an individual bridges two large communities or two small communities. Moreover, it does not consider position changes for each broker depending on the community structure found by the community detection algorithm. To overcome these limits, we developed a robust score that 1) assesses community brokers’ potential reach and control at the meso level for a given partition of the network and 2) accounts for community brokers’ position changes due to different potential community structures.

For each community structure, we located those who have ties in other communities than the ones they are assigned to, the community brokers. We assessed the potential reach and control for each bridge created using three features: the size of the population of the community reached (size), its cohesion (efficiency in connecting to other communities) and whether the bridge is exclusive (exclusivity). First, bridging two small communities of three individuals should have less weight than bridging two communities containing dozens, or hundreds of individuals. Thus, the size of the population reached is quantified by counting the number of people in the community bridged by the community broker. Second, even if an individual bridges a large community, it may not mean much if the community is disconnected or sparse;

the more cohesive the community, the more efficiently can information or products flow from the broker to most community members. To assess the cohesiveness of a community, we used Orman, Labatut and Cherifi (2011)'s distance measure, which calculates the average over all distances between two pairs of nodes that correspond to the shortest path between them. Third, bridge exclusivity matters: if two or more individuals from the same community bridge the same additional community, then the bridge created is redundant and the community broker does not have complete control over the product and information flows between the communities bridged. However, as more and more individuals bridge the same two communities, the impact of a co-broker should diminish. We also considered the community broker's assigned community in terms of population size and cohesion.

Thus, for a community broker B found in a trial run t of the Leiden algorithm that yielded a specific community structure, we developed a *local community broker score* $S_t(B)$, defined in Eq. (1):

$$S_t(B) = \lambda_B + \sum_{c_j \in C(B)} \frac{NP_{c_j}}{cohesion_{c_j}\sqrt{NCB_{c^B, c_j}}}, \; where \; \lambda_B = \frac{NP_{c^B}}{cohesion_{c^B}} \qquad (1)$$

where c_j represents a community j in the set of community $C(B)$ bridged by broker B, and c^B represents broker B's own community. NP represents the number of people in community c_j, $cohesion$ represents the cohesion score of community c_j, and NCB represents the total number of co-brokers from the broker's community c^B to the bridged community c_j (including community broker B). For each community broker B, we consider the community bridged and divide the number of people in that community by the community's cohesion times the square root of the total number of co-brokers. We divide the size of the community by the cohesion score to account for the relative effort required to reach all members of a bridged community. We multiply cohesion by the square root of the number of co-brokers to mitigate the impact of a bridge if other individuals are also bridging the two communities. The square root of NCB is taken to consider the diminishing impact of each additional co-broker, as each additional co-broker has less and less impact on a broker's score. To find the score $S(B)$ for each trial t of the Leiden algorithm, we first sum the result for each community bridged, thus increasing community brokerage as more and more communities are bridged. Second, we add the broker's own community score c^B, as defined by λ_B, in which NP represents the number of people in the broker's community c^B, and $cohesion$ represents c^B's cohesion.

Since the Leiden algorithm may yield various community structures, we decided to make the score robust to the inherent randomness of community partitioning, by **calculating the average community broker score over all trials of the community detection algorithm**. This reduces a broker's score if the broker does not hold a stable broker position in all trials. Also, averaging over all trials implies that a partition that emerge more often has more weight than an outlier partition (although the outlier partitions are still considered). The *global community broker score* $S(B)$ is defined in Eq. (2):

$$S(B) = \frac{1}{N} \sum_{t=1}^{N} S_t(B) \qquad (2)$$

Where $S(B)$ represents the average of all broker scores found for each trial $S_t(B)$ and N represents the number of trials. For this study, we ran the algorithm 50,000 times to make sure that the statistical distribution of the uncovered partitions was stable. For network scientists wanting to leverage the community broker score presented above for their own datasets, a Python package is available online[9].

3.4 UNDERSTANDING COMMUNITY BROKERS REACH AND CONTROL

To better understand the community broker score, we conducted a series of nonparametric tests that compared the score with other brokerage measures and the severity score we developed. Due to the non-normality of the community broker score, we opted for **Kendall's Tau-b correlation** coefficient to assess the degree of similarity between two sets of ranks (whether ordinal or continuous) with the same set of objects (Abdi, 2007). Kendall's Tau-b correlations adjust for ties and are found to be more robust (in terms of gross error sensitivity) and more efficient (having smaller asymptotic variance) than the widely used Spearman rank correlations (Croux and Dehon, 2009). **Kendall's Tau-b correlations** were calculated using the *scipy.stats*[10] Python package and on two samples: 1) all the individuals in the network and 2) only the community broker sample. The statistics are computed on both samples because Kendall Tau-b's loses accuracy in situations with high number of zeros (i.e., the community broker scores). Considering both samples provide a more holistic understanding of how the score relates to other brokerage and social network measures. Also, using only the broker sample allows us to assess specifically how the ranks differ for the community brokers who were flagged.

We looked at how the score related to Jones, Ma and McNally (2019)'s *bridge strength* and *bridge betweenness* presented above, two similar meso-level brokerage measures. We computed Jones, Ma and McNally (2019)'s measures on the criminal entrepreneur network using

[9] Information about the package can be found here https://github.com/Masarah/community_broker_score
[10] https://docs.scipy.org/doc/scipy/reference/stats.html

the *bridge* function in the *networktools*[11] R package. Since the two measures require a predefined community structure to be computed, we used the most common partition found via the Leiden algorithm. This partition was found in 38% of the trials, had a modularity score of 0.926 and 39 groups. Then, we looked whether associations exist between the community broker score and traditional networks measures (*degree centrality* and *Freeman betweenness centrality)*, structural hole measures (e*ffective size, efficiency and constraint)*, age and the severity score mentioned above. All network centrality and structural hole measures were calculated using the Python *NetworkX*[12] package.

4. RESULTS

The network of criminal entrepreneurs studied yielded stable individual positioning in terms of community brokers across variations in community structures. We found 486 unique network partitions of 37, 38, 39 and 40 groups. All network partitions yielded high modularity scores, ranging from 0.923 to 0.926. The average *Adjusted Rand Index* score, considering each partition (also known as community structure) pairs, was 0.91 (std = 0.04) with a minimum of 0.79 and a maximum of one, illustrating the strong similarity between each partition pair found in the network via the Leiden algorithm. Moreover, the proportion of community brokers in the whole network remained stable. These individuals represented, on average, 9.2% of the network population for each trial (std=0.002, min = 9.2 and max = 10.1).

Scoring community brokers requires considering community brokers in all trials of the community detection algorithm. A total of 247 individuals out of the 1,803 criminal

[11] https://rdrr.io/cran/networktools/man/bridge.html
[12] https://networkx.org/

entrepreneurs were, at least once, a community broker in one trial and 1,556 were never brokers. Of these 247 community brokers, 95 (38%) remained community brokers throughout all trials. By adding these 95 community brokers with the 1,556 individuals who are never brokers, we can conclude that 92% of criminal entrepreneurs in the network had a stable meso-level positioning, either as community brokers or non-brokers.

4.1 THE STRENGTH OF THE COMMUNITY BROKER SCORE

The community broker score was calculated for each of the 247 individuals who were found in that position for at least one trial. This section starts by showing how the community broker score behaves using specific examples while the next one presents the results of the community broker score in the network studied.

Measuring Potential Reach and Control: To illustrate how the collective broker score behaves, we selected the most common partition of 39 communities, and its local community brokers score $S_t(B)$ distribution. In this specific partition, the average number of individuals forming a community was 46 (std = 27), with a minimum of 11 and a maximum of 121 individuals. We looked at the meso-level networks for the community broker with the highest score (A) and the two with the lowest score (B and C as there was a tie). Fig. 1 presents the meso-level network for community broker A and Fig. 2 for B and C. In both figures, the community brokers are represented with larger nodes and their bridging ties display a wider width. The communities are tagged based on the numbers given by the Leiden community detection algorithm. As shown in Fig. 1, community broker A belongs to community 10 (cohesion = 3.78, N=64) and bridges four large communities: community 1 (cohesion = 3.33, N=112), 4 (cohesion = 4.21, N=79), 6 (cohesion=3.83, N=70) and 17 (cohesion= 3.20, N=40). Community broker A scores the highest (84.87) in this partition because all communities bridged are large and cohesive and because the broker holds a unique bridge for community four and 17 and shares a bridge to community one and six with only one individual.

[INSERT FIGURE 1 ABOUT HERE]

On the other hand, Fig. 2 shows meso-level network for community brokers with the lowest scores in the partition. Both brokers belong to community 35 (cohesion=1.78, N=16) and share a bridge to community 10 (cohesion = 3.78, N=64). They are both bridging a large community of 64 people, but they belong to a small community (N=16) and since they are sharing the tie, their community broker score decreases.

[INSERT FIGURE 2 ABOUT HERE]

Fig. 3 presents the community structure of the whole network to provide a better representation of where these community brokers fit in the larger social network. The color of the nodes represents their community membership, and their size are scaled based on the community broker score. Community broker A, B and C are also marked in the graph. Why these three brokers are positioned at the extreme of the community broker score distribution is obvious: community broker A is situated in the center of the network, with many bridging ties, while community broker B and C are part of a smallest community and share a bridge to another community.

[INSERT FIGURE 3 ABOUT HERE]

Figure 3 only illustrates one representation of the community structure, while many other structures were found by the Leiden algorithm. Therefore, the score $S_t(B)$ is calculated for each trial and the community broker score $S(B)$ is the average over all trials. The measure is thus not dependent on which partition is chosen arbitrarily but gives more weight to network partitions that are more often uncovered while also considering outlier partitions. To illustrate how the measure penalizes unstable community brokers, Fig. 4 shows the community broker score distribution for the most common partition (a) and once all partitions are considered (b). When considering all partitions, the score distribution shifts to the left, with a higher number of community brokers scoring close to zero because they do not have a stable broker position. The community broker score $S(B)$ found in the criminal entrepreneur's network spans from 0 to 91.31, with a mean of 3.53 (std = 11.49) and a median of zero. The score is compared to other brokerage measures below.

[INSERT FIGURE 4 ABOUT HERE]

4.2 COMMUNITY BROKER SCORE: A UNIQUE MESO-LEVEL MEASURE

How does the community broker measure relate to the meso-level bridging measures developed in Jones, Ma and McNally (2019)? To begin, we calculated the bridge strength and bridge betweenness measures on the most common partition found and their descriptive statistics are presented in Table 2, along with the community broker score $S(B)$.

[INSERT TABLE 2 ABOUT HERE]

The mean number of communities bridged by an individual is 0.15 (std = 0.57), mainly because most individuals do not bridge any communities. The minimum number of communities bridged is zero and the maximum is nine. For the bridge betweenness measure, we find that the minimum score is zero and the maximum is above three million, meaning that one node sits at the shortest path between two nodes from different communities up to three million times, the mean is 41,853 (std=173,677) and the median is zero. The community broker score $S(B)$, on the other hand, spans from 0 to 91.31, with a mean of 3.53 (std = 11.49) and a median of zero.

[INSERT TABLE 3 ABOUT HERE]

As shown in the Table 3, when considering the entire sample, there is a moderate to strong relationship between the community broker score and bridge strength (τ_b = 0.82, p-value<0.000) as well as with bridge betweenness (τ_b= 0.50, p-value<0.000). These relationships, however, weaken when we consider only the community broker sample with τ_b=0.67 (p-value<0.000) for bridge strength and $\tau_b=0.27$ (p-value<0.0000) for bridge betweenness. The higher correlation coefficients found when using the full sample are mainly due to the relatively high frequency of zeros; ultimately only a minority of individuals are brokers between communities.

At first glance, the community broker score and the bridge strength measure may be considered quite similar. However, the community broker score accounts for additional variables (numbers of people reached, cohesion and bridge exclusivity) and various network partitions, and the difference in information grasped by the two score is apparent when only community brokers are selected (τ = 0.67, p-value<0.0000). In other words, accounting for alternative options to reach multiple communities changes the relative importance of community brokers in the sample. Thus, if one wishes to assess brokers' potential control over products and information flow in a network, the community broker score may be most appropriate. In other situations where population size, cohesiveness or co-brokers do not matter as much, or when the community structure is known, bridge strength can be an appropriate measure that is more easily calculated.

Bridge betweenness is, in some ways, a more refined measure than bridge strength, as it considers where a node is positioned within a community, and if it sits at the shortest path between two nodes from different communities. However, the measure is computationally intensive and is closely related to Freeman's betweenness centrality measure ($\tau_b = 0.91$, p-value<0.000 for the whole sample and $\tau_b = 0.89$, p-value <0.000 for the community broker sample). More importantly, bridge betweenness may provide high scores to nodes that do not bridge communities. Consider the simulated network in Fig. 5, with three communities and 18 nodes. The community broker score, as well as the bridge strength and bridge betweenness measures, for five key nodes in the simulated network (A2, B1, B2, B3, C1) are presented in Table 4.

[INSERT FIGURE 5 ABOUT HERE]

[INSERT TABLE 4 ABOUT HERE]

Table 4 shows that A2, B1, B3 and C1 are all community brokers. C1 and B3 as well as B1 and A2 are two pairs that have the same broker score: they bridge the same communities without any co-brokers. First, notice that bridge strength does not differentiate between B3 and B1, as they all bridge only one community, but the community broker score does, giving B3 a slightly higher score because it bridges a community with more individuals. Second, notice that even if B5 is not a broker, it scores higher in terms of bridge betweenness because the node sits on the shortest path of B6, B7, B8 and B9 for both the pink and the grey communities. However, B5 has no direct link with these communities, whereas B1 is directly connected to the grey community and B3 is directly connected to the pink community. In this situation, B5 scores higher in terms of bridge betweenness because it is strategically positioned in the community, but the node is not a community broker per se. To appropriately quantify meso-level power, those positioned *in-between* communities should score higher than those who are not, as measured by the community broker score.

4.3 EXCLUSIVE BROKERAGE POWER AND STRATEGIC POSITIONING

Are criminal entrepreneurs with more reach and control at the meso level also different on other network characteristics? We compared the community broker score with specific traditional network and structural hole measures as well as the severity score. The descriptive statistics for each measure are presented in Table 5.

[INSERT TABLE 5 ABOUT HERE]

As shown in Table 5, individuals in the network have, on average, four connections, with the maximum being 41 and the minimum one. The effective size of their network (non-redundant contacts) is slightly lower at a mean of 2.5. Individuals show wide variations in terms of Freeman betweenness centrality, with a mean of 7,721 (std = 30 922) as well as a minimum and a median of zero and a maximum of 627,617. In terms of attributes, 79% of the sample are men aged 33 (std=9), on average, and with a mean severity score of 2.60 (std=1.74).

Table 6 presents Kendall Tau-b correlations between the community broker score and these measures. As shown in Table 6, none of these measures overlap with the community broker score.

[INSERT TABLE 6 ABOUT HERE]

When considering the whole sample, we find moderate positive relationships between community broker score and degree centrality (τ_b= 0.33, p-value=0.000) as well Freeman betweenness centrality (τ_b = 0.48, p-value=0.000). When considering only the community broker sample, the relationship weakens with τ_b= 0.16 for the former (p-value=0.0005) and ρ= 0.16 for the latter (p-value=0.0000). We find a positive moderate relationship between the community broker score and effective size (τ_b = 0.43, p-value=0.000) as well as a negative relationship with constraint (τ_b= -0.38, p-value=0.000). We find no relationship between the community broker score and efficiency (τ_b= 0.04, p-value=0.0371). When considering only the community broker sample, the relationships with effective size (τ_b = 0.18, p-value=0.0001) and constraint (τ_b= -0.19, p-value=0.0000) weaken, while a weak positive relationship with efficiency (τ_b= 0.17, p-value=0.0000) is found. As for the other attributes, we find that there are no significant correlations between age and community broker score for both samples. In terms of *seriousness* of involvement in criminality, proxied by the severity score variable, we find a weak positive relationship (τ_b = 0.11, p-value = 0.000) when considering the whole sample. However, once we consider only the community broker sample, the relationship disappears (τ_b = 0.02, p-value=0.6226). In terms of gender, a total of 85% of community brokers are men, compared to 79% in the entire sample.

5. DISCUSSION

The findings of this study build up on an increasing number of studies that examine the patterns and properties of illicit networks (Bouchard, 2020; Bright, Koskinen and Malm, 2018; Morselli, 2009; among others). With increased data availability, network scientists have started to look at illicit network at the mesoscopic level (Ouellet, Bouchard and Charette, 2019; Lantz and Hutchison, 2015), also investigating those who bridge communities -the community brokers (Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017; Shaefer et al., 2017). We add to the current literature by arguing that the strong community structures found within a criminal entrepreneurs' network is kept together by community brokers who stand out not simply because they bridge unconnected individuals generally, but because they do so at the meso-level. That is, community brokers allow participants in illicit markets to reach each other, even when they belong to different subgroups or communities. Individuals in large illicit networks even prefer to build indirect ties through brokers, for increased security (Bright, Koskinen and Malm, 2018).

Moreover, community brokers seem to represent a small proportion of individuals in illicit networks. They are few and far between; only 9% of the criminal entrepreneurs connected to multiple communities, a finding that is similar to DellaPosta (2017) who found that bridging ties between different communities of a mafioso network were concentrated among a minority of actors. Schaefer et al. (2017) also found that 14.5% of the incarcerated individuals they studied – 19 out of 131, belonged to multiple communities.

Meso-level brokerage in the sorts of illicit networks we studied, we suspect, requires specific qualities that go beyond what we could measure in our study. For instance, DellaPosta (2017) found that those bridging communities in a mafia tend to be of high status, but also of low status. In our study, those with meso-level power were not found to be particularly violent, or old, nor were they particularly connected, compared to others. This may be explained by the varied set of 1,800 criminal entrepreneurs we studied, individuals who were connected through social, criminal, familial and business relationships, compared to previous studies that have focused on very specific contexts, such as mafioso networks (Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017) and prison networks (Shaefer et al., 2017). Further research could investigate whether community brokerage is related to individual characteristics depending on the types of relationship or the size of network studied. Networks that are already built in with a diverse set of entrepreneurs, like the one we examined, may increase the challenge in connecting to individuals across different communities, thus reducing the number of community brokers uncovered. In this context, being a community broker implies connecting with individuals that belong to different criminal subcultures altogether, or to cross ethnic or racial divides (e.g., Schaefer et al., 2017). In mafia networks, connecting across communities may be a privilege that is only given to a select few (Calderoni et al., 2017; Dellaposta, 2017).

The community brokerage measure has potential practical implications for researchers or practitioners interested in network disruption. Prior research has found that removing brokers from illicit networks can successfully disrupt them (Duxbury and Haynie, 2019; Bright et al., 2017; Bright, 2015; Malm and Bichler, 2011; Morselli and Roy, 2008), at least in the short term (Bright, 2015; Duij et al., 2014). In these settings, brokers are generally identified through the Freeman's betweenness centrality measure (Bright et al., 2017; Duxbury and Haynie, 2019). Community brokers, on the other hand, are identified most of the time with the help of community detection algorithms (Calderoni, Brunetto and Piccardi, 2017; DellaPosta, 2017; Shaefer et al., 2017). Future disruption studies could assess whether the measure of community brokerage we propose is more successful at disrupting illicit networks than existing ones. Given their strategic positioning across communities, their potential to fragment networks is high. However, it is important to note that the community broker score was not conceptualized with network disruption in mind. It focuses on quantifying meso-level brokers reach and control over information and product flows within a network. Hence, there might be a need to develop meso-level scores that specifically optimize for researchers who have different aims.

5.1. Leveraging the Community Broker Score

By focusing on quantifying community brokers' reach and control over products and information flows, the score presented taps into a different meso-level brokerage power than Jones, Ma and McNally (2019)'s measures, as well as other brokerage measures (Freeman betweenness centrality, structural hole measures and bridging centrality). It gives more power to individuals bridging many large and cohesive communities, accounts for a loss in advantage when other individuals bridge the same communities and penalizes those that do not have a stable broker position across multiple instantiations of a community detection algorithm. Moreover, averaging the score over many network partitions strengthen the measure, providing a more accurate estimate of community brokers' meso-level power by minimizing the potential impact of missing data or arbitrarily selecting wrong network partitions.

That said, the results that emerged from this study need replication and extension to different contexts and markets. Although we cannot exclude that community brokerage in alternative data may correlate more strongly with traditional brokerage measures, the community broker score (combined with the Leiden algorithm) gave more weight to individuals who are exclusively and steadily positioned in-between large and cohesive communities and these individuals were not easily uncovered with any other measures, at least in this network. They were the hidden brokers that were challenging to capture with other measures.

Lastly, the score developed can be computed with other community detection algorithms, opening a vast array of potential research line that can focus on mesoscopic brokerage, especially considering the increased data availability, both in licit and illicit network research (Bright, Brewer and Morselli, 2021). It can also be modulated according to specific features of a community, as illustrated in the example below.

5.1.1 Example: Community Brokers and Control over Gun Flows

Community brokers hold power over information or asset flows between communities. From the perspective of illicit product distribution, almost any product need to go through them. Consider the example of firearm distribution. From a meso-level perspective, community brokers are the ones controlling firearm circulation in-between communities. Thus, rather than considering the number of people in a community, one could consider the number of firearms and assess community brokers' potential reach and control over firearm distributions in illicit networks.

As an example, the variable "the number of registered firearms" is used below as a proxy for legal gun access among criminal entrepreneurs (an actual attribute available in the data we received). From a meso-level perspective, the number of legal guns becomes a community feature. To illustrate the argument, we selected a community in the most common partition and investigated its access to legal guns through community brokers. Figure 6 illustrates that community (red) and the surrounding communities connected through it. The darker edges represent brokering edges from this red community to other communities.

[INSERT FIGURE 6 ABOUT HERE]

Based on this network partition, the red community can reach four other communities, three of which include individuals with legal guns. The guns can only be reached through the brokering ties giving community brokers A, B and C power over the gun circulation in their community. The search for legal guns could be a feature of interest and one could wish to score brokers in terms of how efficiently and exclusively they can bridge other communities that have legal guns. In these situations, the local community broker score modulated for gun access for one trial takes the form presented in Eq. (3):

Eq. (3) $$GS(B) = \sum_{c_j \in C_g(B)} \frac{NG_{c_j}}{cohesion_{c_j} \sqrt{NCB_{c^B, c_j}}}$$

$$where\ C_g(B) = \left\{c_j \in C(B) \mid NG_{cj} > 0\right\}$$

Where c_j represents community j in the set of communities with a positive number of guns C_g bridged by broker B and c^B represents the community of broker B. NG represents the number of legal guns in community c_j, $cohesion$ represents community c_j's cohesion score and NCB represents the total number of co-brokers from the brokers' community c^B to community c_j. The formula is similar to the one in Eq. (1), except that instead of calculating the number of people in a community, the number of legal guns is considered. Also, we only consider the communities with a positive number of legal guns and λ_B, representing the brokers' community score is removed, since, in this specific example, we are not looking for legal guns in the broker's community. Eq. (3) is the local broker score modulated for gun access. For a robust global score, one should calculate the average local score over multiple partitions, as shown in Eq. (2).

Given this partition, the local community broker score can be inferred: community broker A score the highest, followed by community broker C and then B. Community broker A scores higher because the person is bridging two communities, the pink community (high cohesion score) with 12 legal guns and the blue community (medium cohesion score) with six legal guns. However, community broker C also bridges the blue community with six legal guns and the green community with 12 legal guns (low cohesion score); the main difference is that the bridge with the green community is shared with community broker B, explaining community broker C's lower score compared to A. Lastly, community broker B scores the lowest because the person is bridging the green community of 12 guns along with community broker C, reducing the person's control over gun flows. All in all, this example illustrates that the score has the flexibility to measure different kinds of meso-level exclusive brokerage power over specific illegal products.

6. LIMITATIONS AND FUTURE DEVELOPMENTS

This study sets the stage for studying community brokers and provides a strategy to measure their brokerage power. Nevertheless, there remain several limitations that needs to be acknowledged. A first limitation relates to the assumption that one holds meso-level power regardless of the type of relationship. The extent to which one holds meso-level brokerage power may depend on the types of relationship developed by community brokers (such as social, criminal or for business types). Negative relationships could also lead to little meso-level brokerage power for a community broker. To overcome this limit, qualitative data on the network studied are needed, an important challenge that network scientists currently face (Bouchard, 2020).

Moreover, the data available for this research did not allow us to investigate traditional outcomes associated with brokerage, such as measures of criminal achievement (Morselli and Tremblay, 2004). Future research could leverage the community broker score with a richer set of attributes, to better assess whether community brokers do indeed reap the benefits of their strategic positions. There is always the possibility that they are more visible than others, less attached to a particular community and as such, more susceptible to arrest (Morselli, 2010). Relatedly, our data do not allow us to verify whether community brokers are aware of their advantageous position. This is especially important considering that social capital can only be beneficial once individuals are aware of their strategic position, and use it to their advantage (Lin, 2002). Qualitative interviews and survey studies on such population of criminal entrepreneurs may shed light on whether they are indeed aware of their social network position, and whether they actively seek to maintain it (Boissevain, 1974; Morselli, 2001).

Finally, the community broker score relies on a large number of community structures (partitions) found through Leiden community detection algorithm. A total of 95 community brokers (out of 247) maintained their status throughout all trials. This means that 152 community brokers had an unstable meso-level positioning, varying according to the outcome of the algorithm. It should be no surprise that after each instantiation, those who are more likely to change position (and sometimes community affiliation) are meso-level brokers, as they are the ones at the edge of communities. Their connections, in another network partition, may end up in the same community as them, thus reducing their brokerage power due to the different community structure found by the algorithm. In addition, using another community detection algorithm could have yielded a different community broker sample, depending on the algorithm's optimization function. That the community broker sample is relatively unstable within one algorithm and given different community detection algorithms illustrate that the meso level of a social structure is not in itself an established fact; it relies on a set of assumptions attached to the particular community detection algorithm selected. In this study, we minimized this relative instability by considering the partition distribution when calculating the global community broker score, thus capturing all potential community brokers. However, this approach may not be relevant to control for differences in community structures found through different community detection algorithms. Each algorithm is built with a specific optimization function and has its theoretical relevance on how to partition the network. In this study, we conceptualized communities as dense clusters uncovered by the Leiden algorithm, and interpreted community brokers as having control over information and product flows between these communities, such as illegal firearms. In sum, expanding the test of the community broker measure to different data and contexts is the next logical step in developing network measures that can appropriately describe the role and behavior of criminal entrepreneurs.

7. CONCLUSION

Community brokers are the glue holding large illicit networks together; how they leverage their strategic meso-level position ought to be further investigated. This study sets the stage for such further investigations. In the large network of criminal entrepreneurs studied, we find that community brokers form a small proportion of the network, about 9%, a proportion similar to other meso-level network brokerage studies (DellaPosta, 2017; Schaefer et al., 2017). To quantify their exclusive meso-level brokerage, we introduced a robust, flexible, and open-source score that quantifies meso-level brokerage power, named *community broker score*. This score was compared to other meso-level brokerage measures, as well as other network measures, through a series of non-parametric tests and pedagogical examples. As differences between these measures were found, we argue that the score developed taps into a different pattern than traditional measures, one that focuses on grasping individual reach and control over information and product flows in illicit networks at the meso-level. The study results also show that community brokers with high meso-level brokerage power may be considered *hidden brokers*, as the community score does not strongly correlate with other brokerage measures, or measures such as age, or severity of involvement in crime. Investigating the emergence of communities, as well as the position of key individuals bridging them, can unravel previously unknown information about the structure of illicit networks and the flow of illegal products within them.

8. REFERENCES

Abdi, H. (2007). The Kendall rank correlation coefficient. Encyclopedia of Measurement and Statistics. Sage, Thousand Oaks, CA. pp.508--510. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.66.9039&rep=rep1&type=pdf

Adler, P. A. (1993). Wheeling and dealing: An ethnography of an upper-level drug dealing and smuggling community. Columbia University Press. https://doi.org/10.7312/adle94604

Bavelas, A. (1948). Some problems of organizational change. J. of Social Issues. 4(3):48--52. https://doi.org/10.1111/j.1540-4560.1948.tb01510.x

Boissevain, J. (1974). Friends of Friends: Networks, Manipulators and Coalitions. Blackwell, Oxford.

Bouchard, M. (2020). Collaboration and Boundaries in Organized Crime: A Network Perspective. *Crime and Justice: An Annual Review of Research*, 49: 425-469. https://doi.org/10.1086/708435

Bouchard, M., and Morselli, C. (2014). Opportunistic structures of organized crime. In Paoli, L. (ed.), The Oxford Handbook of Organized Crime. Oxford Univ. Press, New York, pp.88--302.

Bradshaw, T. K. (2008). The post-place community: Contributions to the debate about the definition of community. Comm. Dev. 39(1):5--16. https://doi.org/10.1080/15575330809489738

Bright, D., Brewer, R., and Morselli, C. (2021). Using social network analysis to study crime: Navigating the challenges of criminal justice records. Social Networks. 66:50--64. https://doi.org/10.1016/j.socnet.2021.01.006

Bright, D. A., Hughes, C. E. and Chalmers, J. (2012). Illuminating dark networks: a social network analysis of an Australian drug trafficking syndicate. Crime Law Soc Change. 57:151–176. https://doi.org/10.1007/s10611-011-9336-z

Bright, D.A., Koskinen, J., and Malm, A. (2018). Illicit network dynamics: The formation and evolution of a drug trafficking network. J. Quant. Crim. 35: 237—258. https://doi.org/10.1007/s10940-018-9379-8

Burt, R. S. (2009). Structural holes: The social structure of competition. Harvard university Press. Cambridge, Massachusetts, and London.

Burt, R. S. (2004). Structural holes and good ideas. Am. J. Soc. 110(2):349--399. https://doi.org/10.1086/421787

Calderoni, F. (2011). Strategic positioning in mafia networks. 3rd An. Illicit Networks W. Équipe de recherche sur la délinquance en réseau. 1--21. http://www.erdr.org/textes/calderoni.pdf

Calderoni, F., Brunetto, D., and Piccardi, C. (2017). Communities in criminal networks: A case study. Social Networks. 48: 116--125. https://doi.org/10.1016/j.socnet.2016.08.003

Croux, C., and Dehon, C. (2010). Influence functions of the Spearman and Kendall correlation measures. Stat. Methods App. 19(4):497--515. https://doi.org/10.1007/s10260-010-0142-z

Dao, V. L., Bothorel, C., and Lenca, P. (2020). Community structure: A comparative evaluation of community detection methods. Network Science. 8(1):1--41. DOI: https://doi.org/10.1017/nws.2019.59

DellaPosta, D. (2017). Network closure and integration in the mid-20th century American mafia. Social Networks. 51:148--157. https://doi.org/10.1016/j.socnet.2016.11.005

Desroches, F. (2007). Research on upper-level drug trafficking: A review. J. of Drug Issues. 37(4): 827--844. https://doi.org/10.1177/002204260703700405.

Duxbury, S. W., and Haynie, D. L. (2019). Criminal network security: An agent-based approach to evaluating network resilience. Criminology. 57(2):314--342. https://doi.org/10.1111/1745-9125.12203

Eck, J. E., and Gersh, J. S. (2000). Drug trafficking as a cottage industry. Crime Prev. Studies. 11: 241--272. https://www.ojp.gov/library/abstracts/drug-trafficking-cottage-industry-illegal-drug-markets-research-prevention-policy

Ffautz, H. W. (1961). Near-group theory and collective behavior: A critical reformulation. Soc. Probs. 9(2):167--174. http://www.jstor.com/stable/799011

Fortunato, S. (2010). Community detection in graphs. Phys.R. 486(3-5):75--174. https://doi.org/10.1016/j.physrep.2009.11.002

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. Sociometry. 35--41.

Fortunato, S., and Barthelemy, M. (2006). Resolution limit in community detection. Proc. Nat. Acad. Sciences. 104(1):36--41. https://doi.org/10.1073/pnas.0605965104

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. Sociological Theory. 201--233.

Gulbahce, N., and Lehmann, S. (2008). The art of community detection. BioEssays. 30(10):934--938. https://doi.org/10.1002/bies.20820

Hu, Y., Chen, H., Zhang, P., Li, M., Di, Z., and Fan, Y. (2008). Comparative definition of community and corresponding identifying algorithm. Phys. Rev. E. 78(2):1-7. https://doi.org/10.1103/physreve.78.026121

Hwang, W., Cho, Y. R., Zhang, A., and Ramanathan, M. (2006). Bridging centrality: identifying bridging nodes in scale-free networks. Pr. 12th ACM SIGKDD Int. Conf. 20--23. ftp://ftp10.us.freebsd.org/pub/tech-reports/2006-05.pdf

Jones, P. J., Ma, R., and McNally, R. J. (2019). Bridge centrality: A network approach to understanding comorbidity. Multi. Beh. R., 1--15. https://doi.org/10.1080/00273171.2019.1614898

Klerks, P. (2001). The Network paradigm applied to criminal organizations. Connections. 3:53--65. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.4720&rep=rep1&type=pdf

Lancichinetti, A., and Fortunato, S. (2009). Community detection algorithms: a comparative analysis. Phys. R. E, 80(5):1--11. https://doi.org/10.1103/PhysRevE.80.056117

Lantz, B., and Hutchison, R. (2015). Co-offender ties and the criminal career: The relationship between co-offender group structure and the individual offender. J Res Crime Delinquency, 52(5): 658--690. https://doi.org/10.1177/0022427815576754

Lin, N. (2002). Social capital: A theory of social structure and action. Cambridge University Press. New York.

Malm, A., and Bichler, G. (2011). Networks of collaborating criminals: Assessing the structural vulnerability of drug markets. J Res Crime Delinquency. *48*(2): 271-297. https://doi.org/10.1177/0022427810391535

Malm, A., Bouchard, M., Decorte, T., Vlaemynck, M., and Wouters, M. (2017). More structural holes, more risk? Network structure and risk perception among marijuana growers. Social Networks. 51:127--134. https://doi.org/10.1016/j.socnet.2017.01.006

Morselli, C. (2010) Assessing vulnerable and strategic positions in a criminal network. J. of Cont. Crim. Just. 26(4): 382--392. https://doi.org/10.1177/1043986210377105

Morselli, C. (2009a). Hells Angels in springtime. Trends in Org. Crime. 12(2): 145-158.

Morselli, C. (2009b). Inside criminal networks. studies of organized crime, Springer, New York, https://doi.org/10.1007/978-0-387-09526-4. https://www.springer.com/gp/book/9780387095257

Morselli, C. (2003). Career opportunities and network-based privileges in the Cosa Nostra. Crime, Law and Social Change. 39(4): 383--418. https://doi.org/10.1023/A:1024020609694

Morselli, C. (2001). Structuring Mr. Nice: Entrepreneurial opportunities and brokerage positioning in the cannabis trade. Crime, Law and Social Change. 35(3): 203--244. https://doi.org/10.1023/A:1011272411727

Morselli, C., Paquet-Clouston, M., and Provost, C. (2017). The independent's edge in an illegal drug distribution setting: Levitt and Venkatesh revisited. Social Networks. 51: 118--126. https://doi.org/10.1016/j.socnet.2017.04.003

Morselli, C., and Roy, J. (2008). Brokerage qualifications in ringing operations. Criminology. 46(1):71--98. https://doi.org/10.1111/j.1745-9125.2008.00103.x

Morselli, C., and Tremblay, P. (2004). Criminal achievement, offender networks and the benefits of low self-control. Criminology, 42(3): 773--804. https://doi.org/10.1111/j.1745-9125.2004.tb00536.x

Natarajan, M. (2006). Understanding the structure of a large heroin distribution network: A quantitative analysis of qualitative data. J. Quant. Crim. 22:171--192. https://doi.org/10.1007/s10940-006-9007-x

Newman, M. E. J. (2006). Modularity and community structure in networks. Proc. Nat. Acad. Sciences. 103(23): 8577--8582. https://doi.org/10.1073/pnas.0601602103

Orman G.K., Labatut V., and Cherifi, H. (2011). Qualitative Comparison of Community Detection Algorithms. In Cherifi H., Zain J.M., El-Qawasmeh E. (eds) Digital Information and Communication Technology and Its Applications. Springer, Berlin. pp 265-279. https://doi.org/10.1007/978-3-642-22027-2_23

Ouellet, M., Bouchard, M., and Charette, Y. (2019). One gang dies, another gains? The network dynamics of criminal group persistence. Criminology. 57(1): 5--33. https://doi.org/10.1111/1745-9125.12194

Palla, G., Derényi, I., Farkas, I., and Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. Nature. 435(7043): 814--818. https://doi.org/10.1038/nature03607

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. An. R. Sociology. 24(1):1--24. https://doi.org/10.1146/annurev.soc.24.1.1

Reuter, P. (1983). Disorganized crime: The economics of the visible hand. MIT Press. Cambridge.

Reuter, P., and Haaga, J. (1989). The Organization of high-level drug markets: An exploratory study. Rand Corp. https://www.rand.org/pubs/notes/N2830.html?src=mobile

Schaefer, D. R., Bouchard, M., Young, J. T., and Kreager, D. A. (2017). Friends in locked places: an investigation of prison inmate network structure. Social networks. 51:88--103. https://doi.org/10.1016/j.socnet.2016.12.006

Shimbel, A. (1953). Structural parameters of communication networks. Bull. Math. Biophysics. 15(4):501--507.

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Scientific Reports. 9(1):1--12. https://doi.org/10.1038/s41598-019-41695-z

Zaitch, D. (2002). Bosses, brokers and helpers. Labour and business relations amongst Colombian cocaine traffickers. Ams. Socio. Tijdschrift. 29(4): 502--529. https://ugp.rug.nl/ast/article/view/30446

Figure 1 – *Meso-Level Network for Community Broker A*

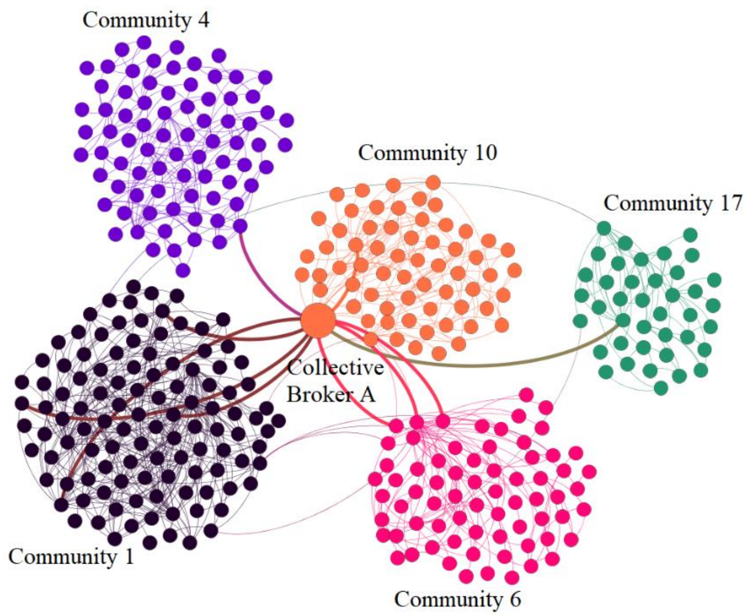

Figure 2 – *Meso-Level Network for Community Brokers B and C*

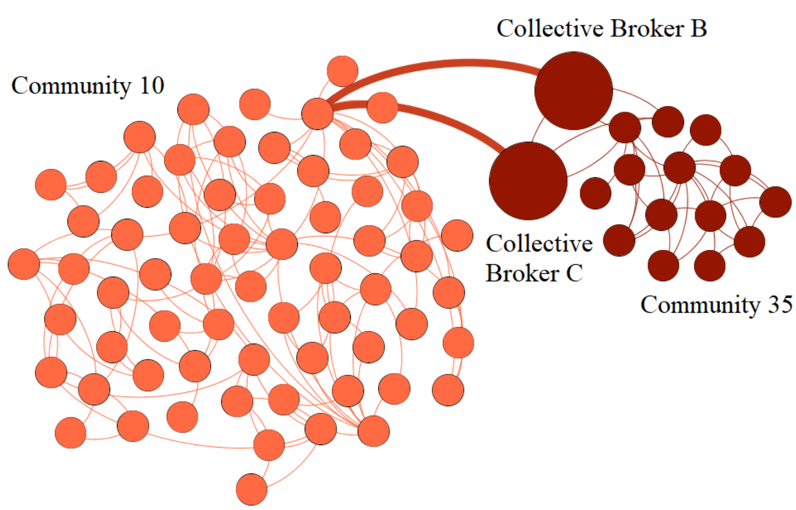

Figure 3 – *Community Structure of the Most Common Partition*

Note. The graph was computed using the HifanHu Proportional and the Noverlap algorithms with Gephi.

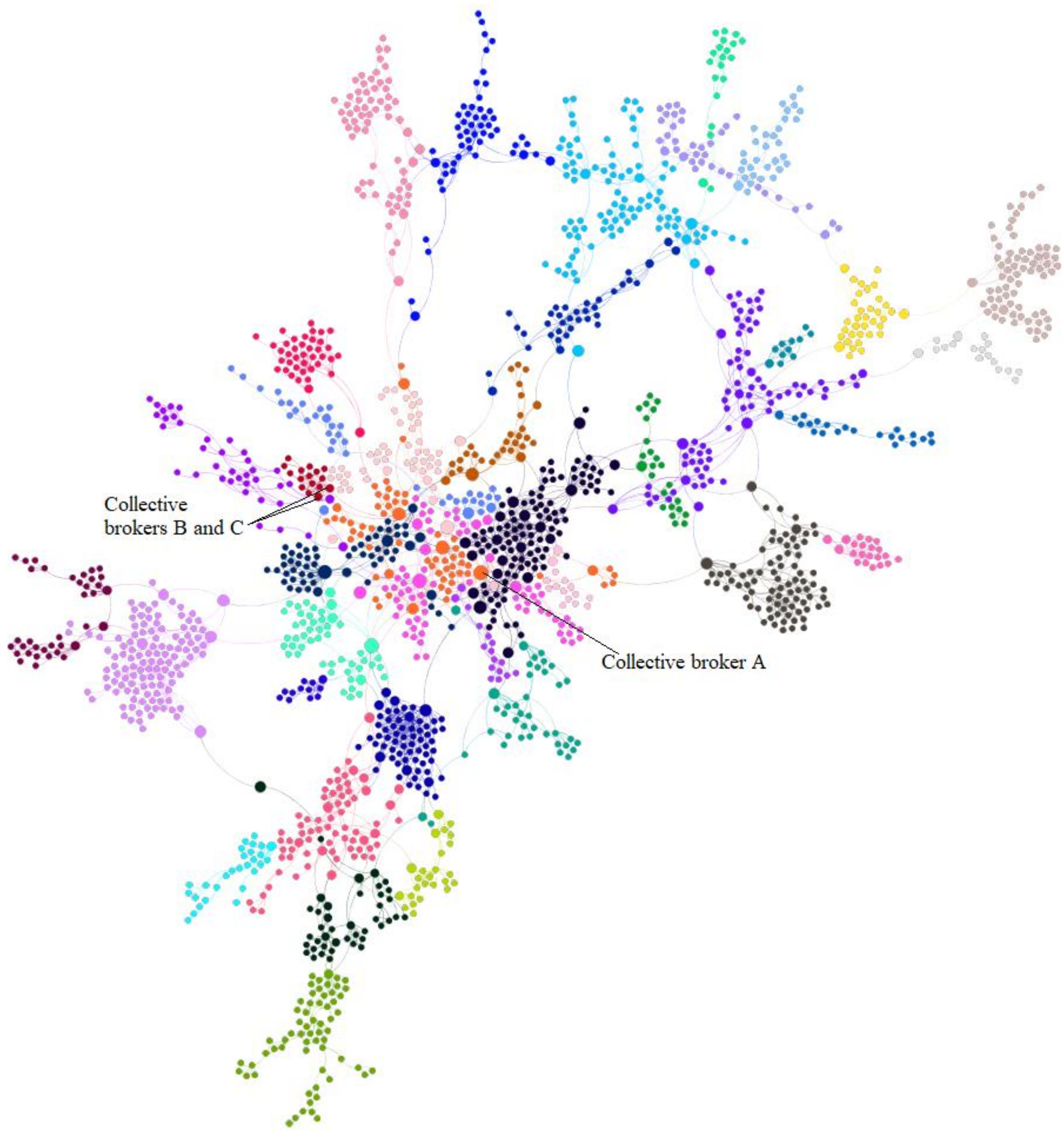

Figure 4a - *Histogram of Community Broker Scores Distribution for the Most Common Partition*

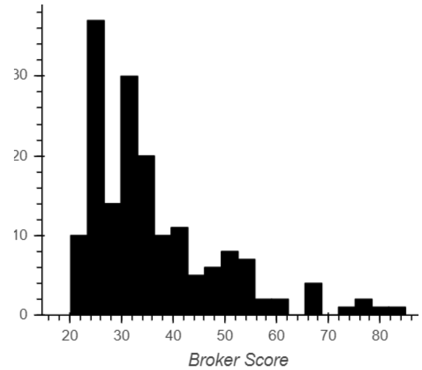

Figure 4b - *Histogram of Community Broker Scores Distribution for all Trials*

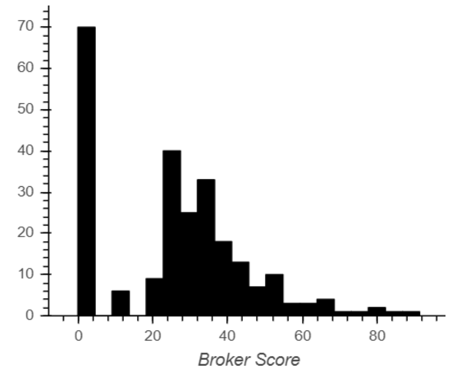

Figure 5 - *Simulated Network*

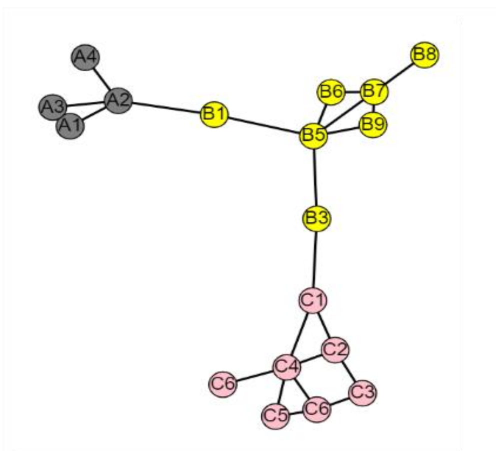

Figure 6 –*Example with Gun Flows*

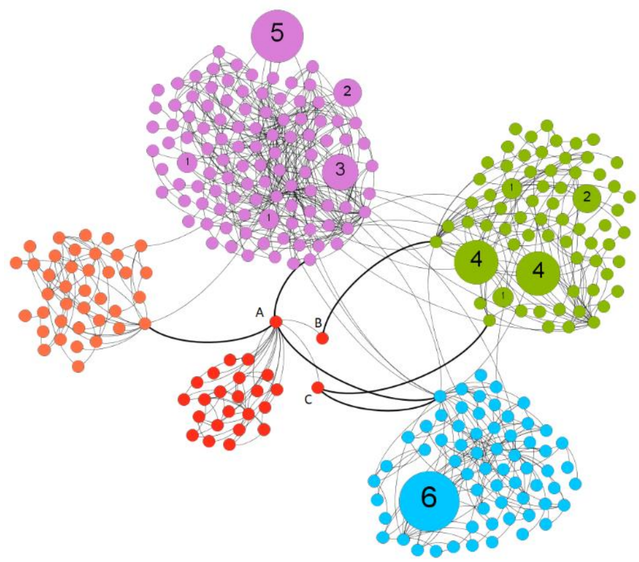

Table 1 – *Personal Attributes*

Attribute	Description
Age	The age of the individual
Gender	Female or male (1=male, 0=female)
FPS (Fingerprint Serial) number	Whether the fingerprints of the individual are recorded in the national criminal database related to the Canadian Police Information Center (CPIC). (1=yes, 0=no)
Engaged in known criminal activity	Whether the individual is known to be actively engaged in criminal activities (1=yes, 0=no)
Flagged armed and dangerous	The individual is flagged as armed and dangerous by the Canadian Police Information Center (CPIC) (1=yes, 0=no)
Flagged violence	Qualitative analysis as to whether the individual had been arrested for a violent offense (1=yes, 0=no)
Weapon prohibition	Whether the individual is prohibited from having a weapon (1=yes, 0=no)
Weapon charge(s)	Whether the individual has weapon charges (1=yes, 0=no)

Table 2 – *Descriptive Statistics Bridging Measures*

	Min	**Max**	**Mean (std)**	**Med**
Bridge strength	0	9	0.15 (0.57)	0
Bridge betweenness	0	3 561 990	41 853 (173 677)	0
Community Broker Score $S(B)$	0	91.31	3.53 (11.49)	0

Table 3 –*Community Broker Score Relationships with Other Meso-Level Brokerage Measures*

	Community Broker Sample N=247	
	Kendall Tau-b (τ_b)	**P-value**
Bridge Strength	0.67	0.0000
Bridge betweenness	0.27	0.0000

Table 4 – *Simulated Network Results on five key nodes*

	Broker	Broker score	Bridge strength	Bridge betweenness
C1	Yes	8.41	1	66
B3	Yes	8.41	1	70
B1	Yes	7.08	1	52
A2	Yes	7.08	1	42
B5	No	0	0	83

Table 5 – *Descriptive Statistics (N=1,803)*

	Min	Max	Mean (std)	Med
Network Measures				
Degree Centrality	1	41	3.91 (3.93)	3
Freeman Betweenness Centrality	0	627,617	7,721(30 922)	0
Burt's Structural Hole Measures				
Effective size	1	36	2.50 (3.33)	1
Efficiency	0.09	1	0.69 (0.28)	0.68
Constraint	0.04	1	0.63 (0.29)	0.60
Attributes				
Gender			M = 79%	
Age	18	81	33 (9)	32
Severity Score	0	6	2.60 (1.74)	3

Table 6 – *Kendall Tau-b correlations between the Community Broker Score and other measures*

	All N=1,556		**Community Broker Sample N=247**	
	Kendall Tau-b (τ_b)	**P-value**	**Kendall Tau (τ_b)**	**P-value**
Network Measures				
Degree Centrality	0.33	0.0000	0.16	0.0005
Freeman Betweenness Centrality	0.48	0.0000	0.29	0.0000
Burt's Structural Hole Measures				
Effective size	0.43	0.0000	0.18	0.0000
Efficiency	0.04	0.0371	0.17	0.0001
Constraint	-0.38	0.0000	-0.19	0.0000
Attributes				
Age	0.02	0.2970	0.06	0.1390
Severity Score	0.11	0.0000	0.02	0.6226