
NOUVELLES POSSIBILITÉS D'ÉTUDE DE LA VULNÉRABILITÉ DES JEUNES À PARTIR DES ARCHIVES ADMINISTRATIVES

S. BALLABIO, A. CARRA, S. CASACCI, D. FERRAZZA, F. VERRECCHIA, A. VITALINI, L. VIVIANO

Istat¹

verrecchia@istat.it

Mots-clés : statistiques expérimentales, archives administratives, ArchIMEDe, vulnérabilité des jeunes, NEET.

1. Résumé

Le papier poursuit deux objectifs principaux. Le premier but est lié à l'évaluation des limites et des possibilités d'utilisation de l'information secondaire pour les statistiques expérimentales à l'appui des politiques locales. Le deuxième but concerne le réexamen des définitions opérationnelles adoptées en tant qu'approximation du concept de vulnérabilité des jeunes, en se concentrant sur le phénomène des NEET (Not in Employment, Education or Training). Les revenus ont été perçus comme une dimension supplémentaire, capable de segmenter la population NEET et de faire lumière sur les zones grises qui les entourent, afin d'identifier d'autres catégories conceptuelles: NIEET (Not Income, in Employment, Education ou Training) et Quasi-NEET - jeunes non engagés dans un parcours de formation et dont le revenu individuel est inférieur au seuil de pauvreté absolue pour une famille unipersonnelle du nord. On utilisera les données du projet ArchIMEDe (ARCHives Intégrées de Microdonnées Economiques et Démographiques) de l'ISTAT, créé avec le but d'étendre l'offre d'information d'intérêt local par la production des microdonnées à partir de sources administratives. Des arbres de décision seront utilisés afin de mieux mettre en évidence la vulnérabilité et les absences contemporaines des réseaux professionnels et de formation, ainsi que des réseaux sociaux et familiaux.

Abstract en anglais

The article focuses on the experimental use of administrative data about residents in Lombardy from the ArchIMEDe project (ARCHives Intégrées de Microdata Economiques et Démographiques) of the Italian Institute of Statistics (ISTAT) and has two main objectives. The first objective is to evaluate the limits and potential use of secondary information for experimental statistics in support of local policies. The second objective is to review the definitions and operational measures of the phenomenon of NEET (Not in Employment, Education or Training), by using income, as an additional dimension to those traditionally considered, able to segment the population and shed light on the grey areas surrounding it, in order to identify other conceptual categories: NIEET (Not Income, in

¹ Il convient de souligner que l'œuvre ne reflète pas nécessairement les vues de l'Institut. Les opinions exprimées dans cet article sont celles des auteurs et n'engagent pas l'institution à laquelle ils appartiennent. L'œuvre est sous la responsabilité conjointe des auteurs. Le paragraphe 1 est attribué à Sara Casacci, le paragraphe 3.1 est attribué à Arianna Carra, le paragraphe 3.2 est attribué à Daniela Ferrazza, le paragraphe 3.3 est attribué à Simona Ballabio.

Employment, Education or Training) and Quasi NEET - young people not in education and whose individual income is below the absolute poverty threshold for an individual household in the North.

Introduction

Les données administratives offrent de vastes possibilités d'exploration et d'expérimentation, puisque les informations qu'elles contiennent - si elles sont dûment traitées - peuvent représenter une valeur ajoutée par rapport au détail territorial, à l'augmentation du champ d'information de la production statistique et à la flexibilité de la "construction des données", qui n'est pas planifiée à l'avance, comme c'est le cas, au contraire, avec les enquêtes statistiques.

Dans ce travail, les données administratives ont été utilisées à titre expérimental dans le but de soutenir les politiques locales dans le domaine de l'emploi, de l'éducation ou de la formation, ce qui présente des caractéristiques emblématiques par rapport aux objectifs fixés. En particulier, l'accent est mis sur les NEET (*Not in employment, Education or Training*). Il s'agit d'un phénomène complexe et multidimensionnel, de nature socio-économique, qui présente une manifestation territoriale très articulée et qui fait objet d'une grande attention au niveau des politiques d'intervention. Ce n'est pas un hasard si elle fait l'objet d'une attention particulière au niveau communautaire : depuis 2013, le programme européen Youth Guarantee (Commission européenne, 2016) finance des mesures dans tous les États membres et semble accélérer la dynamique positive du phénomène. En Europe, entre 2012 et 2016, l'incidence des jeunes NEET a régressé de 13,2 % à 11,5 %. En Italie, les fonds de la *Garantie Jeunesse* sont principalement gérés par les Régions appelées à allouer et à affecter les ressources disponibles au niveau sous-régional. Les décideurs politiques s'appuient sur des données statistiques pour guider leurs choix, mais lorsque ces données - comme dans le cas des NEET, dont l'estimation officielle est dérivée d'enquêtes par sondage - n'atteignent pas le niveau sous-régional, il peut y avoir des inefficacités plutôt que des erreurs de prise de décision, même de nature substantielle.

Dans ce contexte, entre en jeu la discussion sur l'utilisation des données administratives pour atteindre un niveau adéquat de granularité territoriale et de spécificité " qualitative ". En effet, la richesse des données permet de descendre à un niveau territorial plus fin et aussi d'observer la population des jeunes à la loupe et d'identifier et d'analyser les sous-populations présentant des caractéristiques spécifiques d'intérêt. En outre, le caractère multidimensionnel des données administratives permet également d'utiliser des dimensions supplémentaires à celles qui sont prises en compte dans les définitions traditionnelles, ce qui permet de problématiser les mesures opérationnelles et les définitions existantes.

Dans le cas spécifique de l'étude du phénomène NEET, considéré comme un indicateur de la vulnérabilité des jeunes, la définition la plus consolidée a d'abord été reproduite au moyen de données administratives. Par la suite, le revenu a été ajouté comme une dimension supplémentaire de l'analyse, en plus de celles traditionnellement considérées, afin de segmenter la population des NEET et de faire la lumière sur les zones grises qui les entourent, jusqu'à ce que d'autres catégories conceptuelles émergent : NIEETs (*Not Income, in Employment, in Education or Training*) et Quasi-NEET - des jeunes non engagés dans la formation et dont le revenu individuel est inférieur au seuil de pauvreté absolue pour une famille monoparentale du Nord.

Par conséquent, l'inclusion du revenu permet non seulement de préciser davantage la population de NEET, mais aussi d'intercepter une sous-population (Quasi-NEET) caractérisée, également, par un degré élevé d'exposition potentielle à la vulnérabilité sociale et matérielle. Il s'agit d'un groupe émergent et potentiellement en expansion, qui semble être particulièrement lié à la nouvelle dynamique du marché du travail et aux nouveaux horizons d'emploi du travail secondaire ou

intermittent. La tendance dans le temps de ces dynamiques semble être bien appréhendée par une dimension réactive telle que celle du revenu, contrairement à d'autres dimensions (plus liées à la réglementation contractuelle) qui détectent les changements avec un certain décalage dans le temps.

La richesse des données administratives a également permis de mettre en relation l'individu et son contexte familial - en tenant compte également du revenu familial équivalent par habitant - mettant ainsi en évidence un autre aspect de l'étude du phénomène, qui est représenté par la présence éventuelle d'un réseau de protection que le jeune peut utiliser dans sa famille (selon le registre de la population).

D'une manière générale, tant sur le plan de la définition que sur celui de l'identification des risques de vulnérabilité mineure ou majeure, l'accent mis sur le revenu nous permet de considérer les autres piliers sur lesquels repose l'État-providence italien, représentés par le réseau familial et le réseau de protection sociale. Dans le cadre européen, s'il est vrai qu'il y a eu contamination des différents types de *welfare*, il est vrai aussi que des spécificités ont été préservées qui doivent nécessairement être prises en compte à la fois d'un point de vue descriptif de phénomènes tels que ceux considérés ici et, surtout, du point de vue de la mise en œuvre de politiques efficaces. En effet, le fait d'être NEET peut prendre différentes formes, et surtout exposer à différents risques de vulnérabilité, selon le type de régime d'aide sociale dans lequel se trouve le jeune.

L'objectif du travail est d'exposer plus en détail la vulnérabilité des jeunes et donc les interventions nécessaires de la politique locale, en essayant de se concentrer sur les domaines de l'absence contemporaine de réseaux d'emploi et de formation, mais aussi de la protection sociale et de la famille. Les analyses ont été réalisées sur des jeunes de 15 à 29 ans résidant en Lombardie, l'un des quatre moteurs de l'Europe et l'une des régions les plus peuplées d'Europe. Les données sont celles du projet ArchIMEDe (*ARCHivio Integrato di Microdati Economici e Demografici*), né avec l'intention d'élargir l'offre d'information de l'ISTAT par la production de microdonnées, à partir de sources administratives.

2. Données du projet ArchIMEDe : la valeur ajoutée de l'intégration des sources administratives

Au cours des dernières décennies, les instituts nationaux de statistique ont encouragé l'utilisation systématique de sources administratives pour la production de statistiques, également à la suite de l'augmentation de la demande de nouvelles informations de la part des utilisateurs, institutionnels et autres. Les sources administratives présentent une richesse considérable d'informations et permettent de produire des données toujours plus détaillées sur le territoire, réduisant ainsi la *statistical burden* et les coûts.

L'exploitation du contenu de l'information provenant de sources administratives est l'un des principaux aspects du projet ArchIMEDe, mis en place en 2013 par l'Institut national de statistique (ISTAT). Le projet vise à élargir l'offre d'information de l'ISTAT en créant des collections de données élémentaires (longitudinales et transversales) sur les familles et les individus, à mettre à la disposition de certains utilisateurs, pour la recherche sociale et économique, la planification spatiale et sectorielle, l'évaluation des politiques publiques au niveau national, régional et local (Garofalo, 2014).

Le travail décrit dans cet article est basé sur l'une des collections de microdonnées produites dans le cadre du projet : la base de données « Conditions socio-économiques des familles – ArchIMEDe », construite avec le but d'être un instrument de connaissance des aspects socio-démographiques, professionnels et économiques des individus et des familles.

Afin de définir le contenu informatif de la base de données, la revue de la littérature sur le sujet, ainsi que l'analyse des informations disponibles dans les archives administratives disponibles, a permis d'identifier quatre dimensions considérées comme significatives, utiles pour la description des

conditions socio-économiques des individus et des familles et pour la possibilité de déterminer situations de vulnérabilité : 1) information démographique et structure familiale, 2) participation aux parcours éducatifs et niveau d'éducation, 3) participation au marché du travail, 4) revenu.

Le cadre théorique adopté pour la spécification des dimensions se réfère à des approches de type "asset-based" (Alwang, Siegel et Jorgensen, 2001) qui, bien que fondées sur une terminologie économique, ont été développées dans de nombreuses disciplines, tant sociologiques et anthropologiques (Moser, 1998 ; Bebbington, 1999), que dans les sciences de l'environnement (Reardon et Vosti, 1995). En littérature, la perspective économique initiale a été développée dans le cadre de l'*entitlement* approche, introduit par Amartya Sen, et à la fin des années 1980 cette approche a été étendue pour inclure le capital social et d'autres formes de biens immatériels.

En général, dans les analyses fondées sur les *asset*, les ménages à revenu moyen-élevé et autres ressources sont considérés comme plus résistants aux pertes potentielles de bien-être causées par des événements défavorables (Moser et Holland, 1997, Freyssinet, 2009, Morrone *et al.*, 2011).

Dans ArchIMEDe, les dimensions identifiées ont été considérées comme des *asset* et, en particulier, sont reconnaissables dans les groupes d'informations suivants :

- 1) **Renseignements démographiques et structure familiales** : permettent de caractériser les individus en fonction des principales variables démographiques (sexe, âge, état civil, citoyenneté) et, d'autre part, d'articuler la composition des familles en termes de nombre et de type. La qualité des variables démographiques et du type et du nombre de familles dépend strictement de la qualité de l'information contenue dans les listes des registres de la population dont elle est tirée et peut être affectée par l'absence de régularisation des situations familiales modifiées ou par le comportement « opportuniste » des individus (par exemple, les résidences fictives qui ne permettent pas de définir correctement les unités familiales).
- 2) **Éducation** : les archives administratives permettent d'identifier les étudiants inscrits dans les écoles et universités italiennes basées en Italie, alors que l'information sur la formation professionnelle n'est actuellement couverte par aucune source. Les informations relatives à la qualification des membres de la famille proviennent du recensement de la population de 2011 et des mises à jour ultérieures tirées des données Miur.
- 3) **Travail** : l'information disponible est liée à la condition d'emploi prévalant en octobre et à l'intensité de travail² pendant l'année. Les sources administratives permettent d'identifier les signes d'emploi régulier dans les entreprises, organismes, etc. situés sur le territoire.
- 4) **Revenu** : La base de données comprend des informations sur le revenu brut perçu par les membres et les ménages au cours de l'année de référence, obtenu sous la forme de la somme des différentes composantes pouvant être retracées à partir des sources. Il convient de noter que les variables se référant à chaque composante du revenu correspondent à une définition d'un type "administratif" imprégné de concepts d'impôt et de salaire, associés à des sous-populations spécifiques de contributeurs et/ou de bénéficiaires de revenus certifiés.

Une fois identifié les dimensions, les sources administratives et de recensement déjà acquises par l'Istat ou les produits réalisés par l'Institut dans le cadre d'autres processus ont été sélectionnés (**tableau 1**).

Puisque l'utilité des données administratives dépend de leurs définitions et de leur couverture, de l'exactitude avec laquelle elles sont documentées et traitées et de la rapidité avec laquelle elles sont

² L'intensité de la main-d'œuvre est une mesure quantitative de la participation au marché du travail, qui met en évidence les situations présentant un risque accru d'exclusion sociale et les effets d'une faible intensité de main-d'œuvre.

mises à disposition (Statistique Canada, 2010), chaque source administrative, avant d'être utilisée, est soumise à certaines procédures visant à sa validation :

- analyse des métadonnées, acquisition d'informations concernant la législation et la stabilité des procédures régissant la production de la source administrative ;
- contrôle des valeurs manquantes ou anormales ;
- contrôles de stabilité temporelle sur les nombres et sur la distribution des variables d'intérêt.

Après ces contrôles initiaux, l'intégration³ des sources et la création de variables dérivées sont effectuées. L'intégration est une caractéristique spécifique de l'utilisation des données administratives, car elles ne sont pas collectées à des fins statistiques et n'observent donc pas toutes les variables d'intérêt et se réfèrent à une population qui ne couvre généralement qu'une partie de la population *target*. En outre, elle vise à harmoniser les données sous un commun dénominateur. Néanmoins, certains problèmes liés à la qualité des sources *d'input* (par exemple, erreurs de couverture, présence de valeurs manquantes) et à l'intégration (tout d'abord les incohérences entre les informations provenant de différentes sources) sont toujours présents et la comparabilité avec les enquêtes, tant du point de vue de la définition que des résultats, n'est pas garantie.

Malgré ces limites, la collecte de microdonnées répond, au moins en partie, à la demande d'informations statistiques territoriales, garantissant une grande souplesse d'analyse, entendue comme la possibilité pour l'utilisateur d'étudier - en fonction de besoins spécifiques - les conditions des groupes de population et d'examiner, par exemple, le lien entre les différents éléments de vulnérabilité, en combinant le niveau individuel et le niveau familial.

Tableau 1 - Archives administratives, collectes de données d'input et principaux contenus informatifs utilisés

Source	Organisme propriétaire	Information utilisée
Le registre de population des municipalités	<i>Municipalités</i>	Données démographiques et structure familiale
Recensement de la population	<i>Istat</i>	Instruction
Registre des étudiants et des étudiants universitaires	<i>Ministero dell'istruzione, dell'università e della ricerca</i>	Instruction
Précarité de l'emploi (a)	<i>Istat</i>	Travail
Archive des déclarations des revenus fiscaux	<i>Ministero dell'economia e delle finanze</i>	Revenu
Dossiers des retraités	<i>Istituto nazionale di previdenza sociale</i>	Revenu
Archive des prestations non liées à la retraite	<i>Istituto nazionale di previdenza sociale</i>	Revenu
Archive des déclarations de contribution d'Uniemens	<i>Istituto nazionale di previdenza sociale</i>	Revenu

(a) **Note** : a) Il s'agit d'une collection de microdonnées visant à identifier les personnes régulièrement employées sur le territoire en fonction du degré de stabilité professionnelle et des principales caractéristiques du travail, réalisées par l'intégration des archives administratives suivantes : *ASIA - DB Occupazione, Gestione Dipendenti Pubblici (Ex INPDAP), Archivio Autonomi Agricoltura, Archivio Lavoratori Domestici*.

3. Catégories conceptuelles

³ L'intégration est garantie par le SIM - Système Intégré de Microdonnées- par l'attribution de codes d'identification uniques valables dans le temps (Ambroselli, 2015).

La flexibilité et le niveau de détail typique des microdonnées administratives permettent des études, des agrégations et des analyses exploratoires qui peuvent déterminer des résultats importants tant du point de vue substantiel que du progrès théorique des définitions opérationnelles.

Tout d'abord, avec les données administratives, il est possible de " reproduire " des indicateurs définis dans d'autres contextes, par exemple dans les enquêtes par sondage. En ce sens, à partir de la définition la plus courante du phénomène, qui définit les **NEET** comme " les jeunes de 15 à 29 ans qui n'étudient pas ou ne travaillent pas ", on peut observer que pour chaque dimension considérée, il existe une source d'information correspondante et analogue dans ArchIMEDe :

- 1) "Les NEET sont des jeunes de 15 à 29 ans. Source : Listes des registres de la population municipaux qui présentent les caractéristiques personnelles de tous les membres des ménages résidents, ainsi que leur taille et leur type, et à partir desquelles on peut obtenir la sous-population des jeunes d'intérêt ;
- 2) "qui n'étudient pas". Source : Registres d'étudiants MIUR qui permettent d'isoler les élèves inscrits⁴ dans les écoles secondaires ou les universités et, par conséquent, de les enlever.
- 3) "et ne travaillent pas. Source : ASIA-Emploi, positions d'assurance Ex INPDAP, travailleurs indépendants dans l'agriculture et travailleurs domestiques qui permettent d'identifier et, par conséquent, d'exclure du calcul ceux qui montrent des signes d'activité sur le marché du travail.

Par conséquent, pour l'estimation des NEET à partir de sources secondaires, un chemin de filtres progressifs appliqués à la base de données initiale est tracé, ce qui fait ressortir la catégorie d'intérêt.

Deuxièmement, la richesse et l'articulation des données d'archives - résultat de l'intégration des sources administratives - offre la possibilité de sélectionner des sous-populations spécifiques grâce à l'utilisation des multiples dimensions/variables mises à disposition. Les Indicateurs ainsi construits peuvent être de meilleurs supports et avoir une portée heuristique et théorique différente. A travers le projet ArchIMEDe, en effet, ils sont disponibles données qui permettent de considérer de multiples dimensions qui jettent un nouvel éclairage sur l'observation des caractéristiques et dynamiques sociales de la classe des jeunes considérés (15-29 ans).

Concrètement, l'inclusion de la variable du revenu individuel, en combinaison avec les autres dimensions qui décrivent le phénomène des NEET, permet d'identifier une nouvelle catégorie conceptuelle : les **Quasi-NEET**. Il s'agit d'un sous-ensemble limité de sujets qui, sans bénéficier d'un réseau de protection, de formation ou de travail, bénéficient d'un soutien financier qui pourrait être lié à un minimal réseau social (par es. les pensions d'invalidité civile, les pensions sociales, les allocations familiales, les allocations de maternité municipales et d'État, etc. ; les allocations périodiques reçues par le conjoint ou l'ex-conjoint.). La possibilité pour les données administratives de mettre en évidence les revenus individuels, également autres que les revenus du travail a fait apparaître un groupe de sujets qui, en réalité, pour des raisons différentes, l'étaient dans ce cas. En fait, en plus du revenu des allocations et des pensions (et des erreurs possibles), il existe, par exemple, de nouvelles formes de micro-emplois qui ne sont pas immédiatement associés aux employeurs. La catégorie émergente des Quasi-NEET semble être en partie liée à la nouvelle dynamique du marché du travail et aux nouveaux horizons d'emploi du travail secondaire ou intermittent.

Après tout, cette dernière catégorie chevauche les catégories (ex)NEET et jeunes travailleurs (**figure 1**) et comprend à la fois le sous-groupe (ex)NEET, qui montre des signes de faible revenu individuel, et le sous-groupe des jeunes travailleurs à revenu individuel qui ne sont pas en mesure de garantir l'indépendance économique⁵.

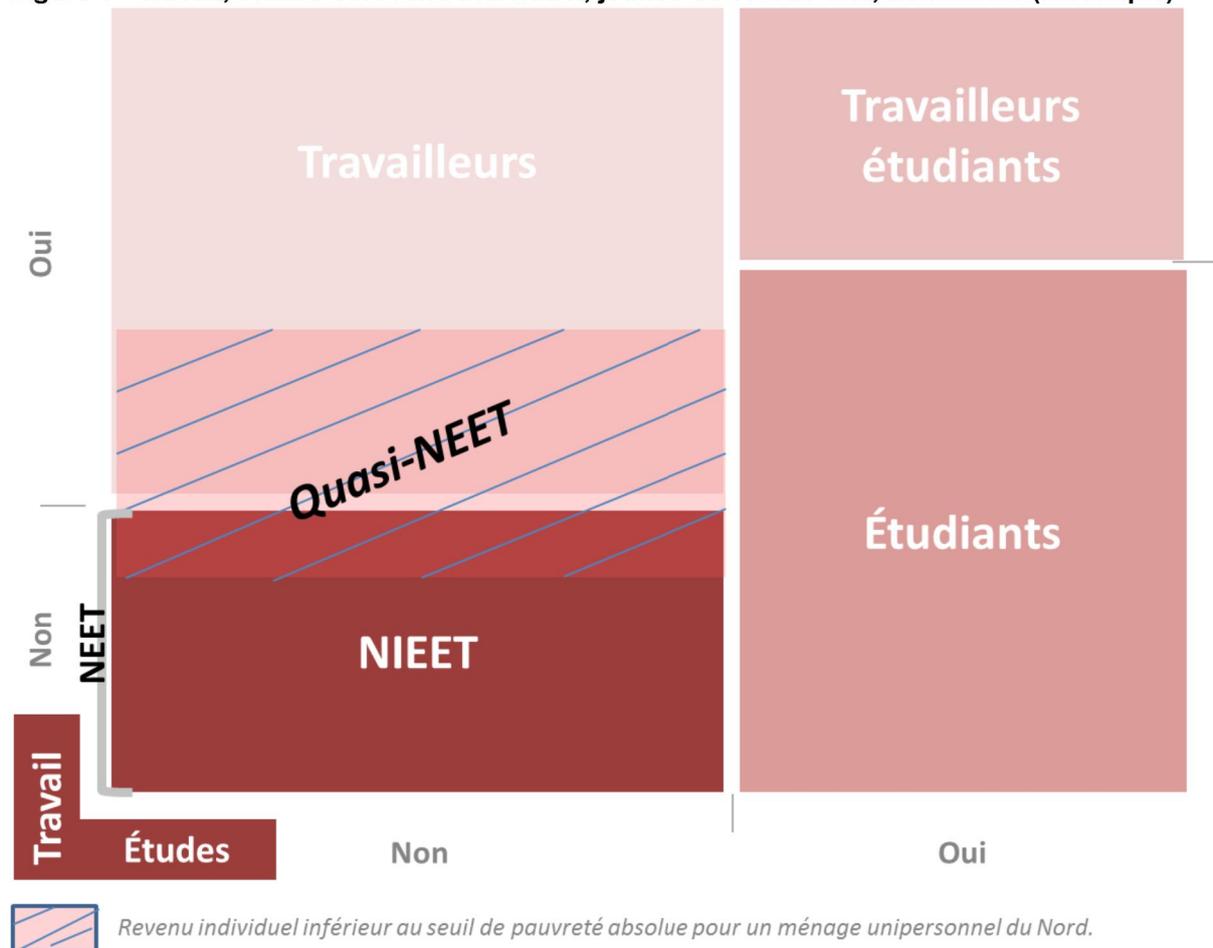
⁴ ArchIMEDe, à ce jour, n'inclut pas les données régionales sur la formation professionnelle.

⁵ Plus précisément, il était considéré comme le seuil de pauvreté absolue pour un ménage unipersonnel dans le nord de l'Italie.

De plus, la prise en compte de la dimension revenu ne peut qu'améliorer la capacité à saisir l'évolution du phénomène dans le temps, puisqu'il s'agit d'une dimension plus réactive par rapport à d'autres qui détectent les changements avec un certain retard dans le temps, puisqu'elles sont liées à la réglementation contractuelle.

Dans ce contexte, pour la conception d'outils d'intervention adéquats, deux nouvelles populations *target* (par rapport à celles des NEET) sont identifiées : **NIEET** (*Not Income Employment Education or Training*), un sous-ensemble de NEET ; et les Quasi-NEET, caractérisés, bien que de différentes manières, par un potentiel degré élevé de vulnérabilité sociale et matérielle.

Figure 1 - Travail, études et revenu individuel, jeunes de 15 à 29 ans, Lombardie (mosaïque)



Note: Ballabio *et al.*, 2017

4. Analyse et résultats

1.1. Structure et dynamique des jeunes en Lombardie

La classification de la population résidente par sexe, groupe d'âge et citoyenneté effectuée sur la base des données et des résultats officiels des derniers recensements a été d'une utilisation exploratoire, car les recensements permettent d'observer les changements survenus, au cours d'une décennie, dans la population de référence, tant en termes de quantité que de composition en fonction des principales variables sociodémographiques.

A la date de référence du recensement de 2011, la population résidente en Italie s'élevait à un peu moins de 60 millions d'individus, dont 16% étaient concentrés en Lombardie. Avec près de 10 millions d'habitants, la région était déjà l'une des plus peuplées d'Europe.

A la même date, les jeunes de 15 à 29 ans représentaient environ 1,4 million d'unités et, dans l'ensemble, 14% de la population lombarde. Si l'on examine les données selon la citoyenneté, il est également clair que l'incidence des étrangers était plus élevée pour le groupe d'âge des 15-29 ans (15%) que la valeur correspondante pour la population lombarde en général (un peu moins de 10%).

Compte tenu de la dynamique, par rapport au recensement de 2001, la population résidente de la Lombardie a augmenté de près de 672 mille unités au cours de la décennie. En supposant que l'augmentation globale est égale à 100, on peut constater que celle-ci est principalement composée de citoyens étrangers (93%) et d'une petite partie seulement (7%) des personnes de nationalité italienne. En ce qui concerne uniquement la partie étrangère de la population, on peut observer que l'augmentation du segment féminin est plus marquée (environ +215% par rapport à 2001) que celle de la partie masculine (+179%).

La dynamique de la classe des jeunes confirme également la tendance esquissée et donne matière à réflexion : sur la décennie considérée, les 15-29 ans de Lombardie ont chuté de 13%. Toutefois, en observant les données sur la base de la citoyenneté, face à une baisse du nombre d'Italiens qui dépasse 22%, on constate une augmentation constante du nombre de jeunes étrangers (+167%). Ces dynamiques opposées, dont on peut penser qu'elles dépendent de multiples facteurs culturels et migratoires, ont, en l'espace d'une décennie seulement, profondément modifié, en termes de citoyenneté, la structure de la population appartenant à la classe des jeunes âgés de 15 à 29 ans. En effet, à partir d'une part de 5 % en 2001, les jeunes étrangers arrivent en 2011 pour constituer plus de 15 % de la population du groupe d'âge correspondant (**tableau 2**).

Compte tenu de ces dernières dynamiques observées - une diminution considérable de la population des jeunes - la population italienne n'est que partiellement compensée par une croissance des résidents de nationalité étrangère appartenant au même groupe d'âge - les analyses qui suivront devront nécessairement prendre en compte la différenciation par nationalité afin d'identifier les clés significatives d'interprétation du phénomène.

Tableau 2 - Population résidente selon le sexe, le groupe d'âge et la citoyenneté, Lombardie, 2001, 2011 (valeurs absolues et pourcentages)

	2001			2011		
	Homme	Femme	Total	Homme	Femme	Total
(#)	4 373 289	4 659 265	9 032 554	4 712 503	4 991 648	9 704 151
Étrangers (%)	4	3	4	10	10	10
Jeunes de 15-29 ans (#)	812 552	778 017	1 590 569	703 613	678 602	1 382 215
Jeunes étrangers de 15-29 ans (%)	5	5	5	14	16	15

Source : Analyse des données ISTAT (recensement de la population de 2001, 2011).

1.2. Vulnérabilité des jeunes en Lombardie et modèle familial

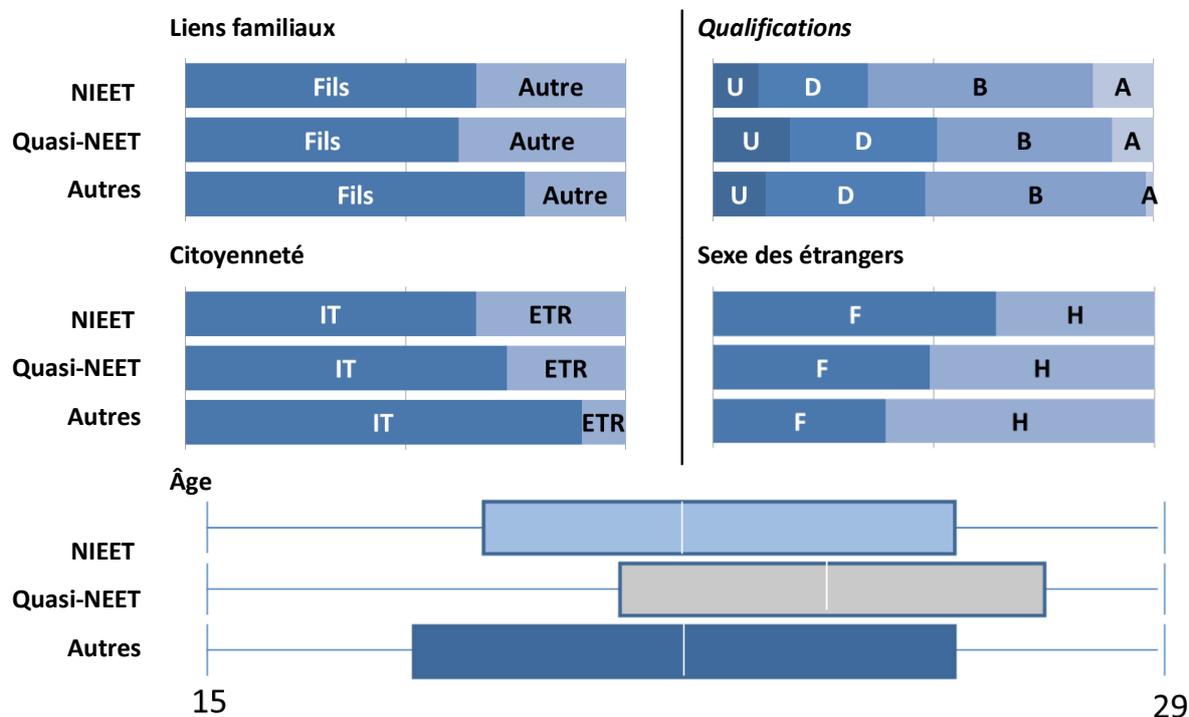
En utilisant les informations contenues dans la base de données, il est possible d'obtenir une description très détaillée des catégories NIEET et Quasi-NEET, en les différenciant du reste de la population âgée de 15 à 29 ans qui, en termes de risque de vulnérabilité, peut être considéré comme un tout. Les variables suivantes ont été prises en compte (**figure 2**) : relations familiales ; citoyenneté ; sexe (des résidents de citoyenneté étrangère seulement) ; qualification ; âge.

En ce qui concerne les relations familiales, dans tous les groupes, la majorité des jeunes vivent dans leur ménage d'origine. En particulier, il ne semble pas y avoir de différence substantielle entre NIEET et Quasi-NEET en ce qui concerne cette caractéristique.

Des différences entre le NIEET et le Quasi-NEET sont observées en ce qui concerne l'âge et le niveau d'instruction : les Quasi NEET ont tendance à être plus âgés et à avoir un niveau d'instruction plus élevé.

Si l'on considère la citoyenneté, par rapport aux NIEET, les Quasi-NEET sont dans une mesure plus grande des citoyens italiens. Si l'on se concentre uniquement sur le sexe des jeunes étrangers, il apparaît que les filles sont plus vulnérables que leurs pairs.

Figure 2 - NEET, Quasi-NEET et autres, selon les liens familiaux, la citoyenneté, le sexe des étrangers, les qualifications, l'âge. Lombardie, 2012



Note: Cfr. Ballabio *et al.*, 2017; Liens familiaux - Fils : jeune vivant avec au moins un des parents ; autre : autre type de relation avec le titulaire de la carte personnelle du registre de la population; Titres - U : Université ; D : Diplôme ; B : jusqu'au brevet inclus ; A : Aucun titre ; Sexe des étrangers- F : Femmes ; M : Hommes ; Citoyenneté - IT : Italiens ; ETR : Etrangers ; Groupes - NIEET : NEET sans revenu ; Autres : 15-29 ans non NIEET et non Quasi-NEET.

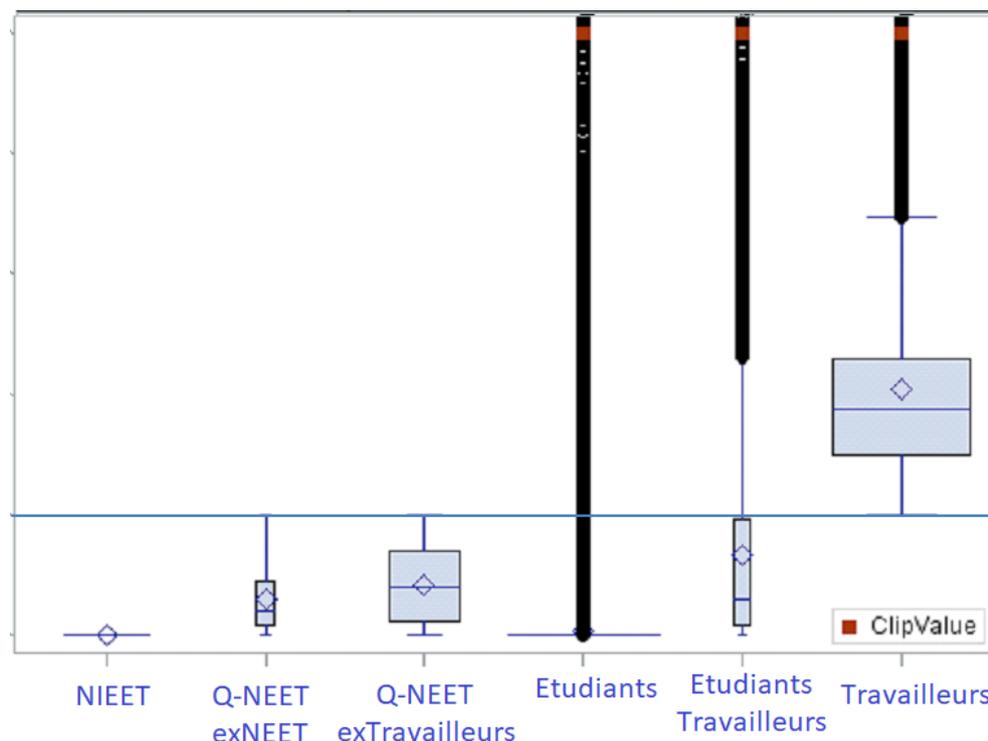
L'utilisation de la variable revenu dans la définition opérationnelle des NEET, en plus de l'identification de nouvelles catégories conceptuelles, démontre sa valeur heuristique également lorsqu'il s'agit d'approfondir les caractéristiques des nouvelles catégories identifiées, pour observer - par exemple - comment le revenu est distribué en leur sein et si et comment il fournit de nouvelles informations substantielles à prendre en compte dans l'élaboration et/ou l'affinement des politiques relatives.

De plus, en exploitant la possibilité de disposer d'informations sur le revenu à différents niveaux de référence (individuel et familial), il est possible de reconstruire une image plus complète, qui réussit à mettre la personne en relation avec son contexte familial. La relation entre l'individu et la famille est toujours importante si nous voulons préparer l'analyse d'un phénomène socio-économique, elle l'est

encore plus si nous voulons approfondir la réflexion sur les questions de *welfare* dans lesquelles, comme on le sait, le rôle de la famille devient connotant. L'un des éléments qui distingue les régimes de protection sociale de certains pays méditerranéens (Italie, Espagne, Grèce, Portugal) consiste dans le rôle spécifique de la famille et de la solidarité parentale tout au long de la vie de l'individu⁶ (Ferrera, 2006). Considérant le cas italien comme une variante spécifique du modèle de *welfare* conservateur-corporatiste (Esping-Andersen, 1990), l'étude de la répartition des revenus dans une clé individuelle et familiale offre la possibilité de réfléchir davantage sur le degré de vulnérabilité des jeunes dont ils veulent s'occuper.

La représentation par box-plot permet d'observer synthétiquement les principales caractéristiques des distributions des revenus individuels au sein des différentes catégories d'analyse identifiées, en gardant dans ce cas toutes les sous-populations identifiées : NIEET, Q-NEET exNEET, Q-NEET exTravailleurs, Etudiants et Etudiants Travailleurs (figure 3). Dans la figure, le seuil de vulnérabilité fixé à 10 mille euros⁷ a été mis en évidence et à gauche il y a les catégories sans aucune forme de protection (en termes de système de formation et de travail) et à droite celles qui sont incluses dans au moins un des deux systèmes. Si l'on considère le revenu individuel, toutes les catégories, à l'exception des Travailleurs, se situent totalement en dessous du seuil de vulnérabilité, avec des valeurs proches de zéro pour les Étudiants et les NIEET. Dans d'autres cas, les Q-NEET exTravailleurs et les Travailleurs montrent des distributions plus équilibrées et symétriques, tandis que les Étudiants-Travailleurs et les Q-NEET exNEET montrent des distributions plus asymétriques vers des valeurs de distribution élevées, qui restent cependant, pour ce dernier, très proches du seuil de 10.000 euros.

Figure 3 - Jeunes de 15 à 29 ans, système éducation-travail et le revenu, Lombardie, 2012 (Box plot)

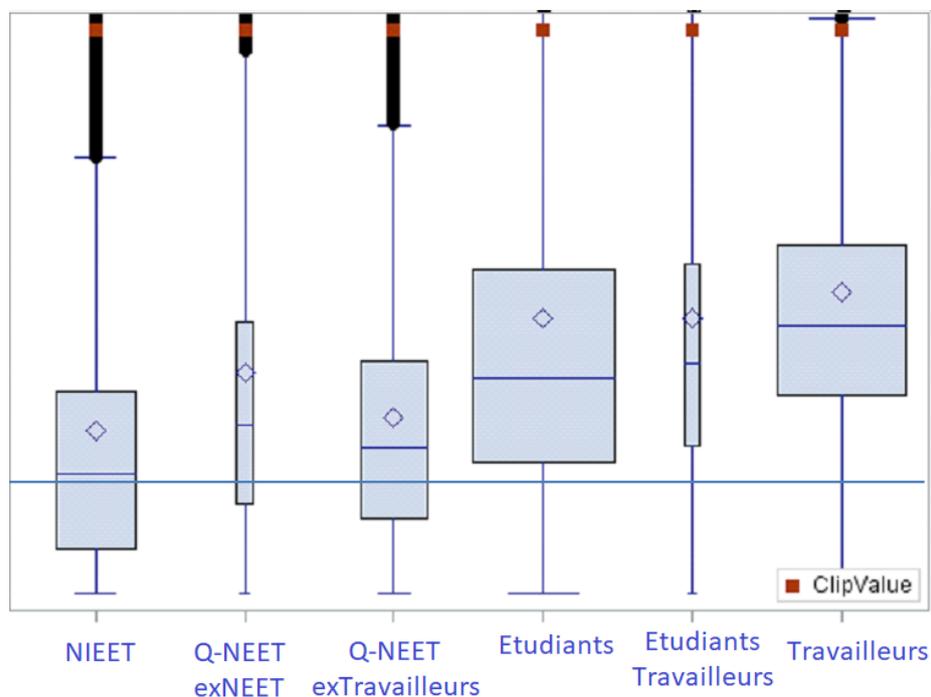


⁶ Les autres caractéristiques saillantes sont la réglementation dualiste du marché du travail (qui divise clairement entre les individus hautement protégés et les individus mal protégés), le caractère universaliste du système national de santé et un degré élevé de particularisme.

⁷ Le seuil de pauvreté absolue (cf. ISTAT - Pauvreté en Italie, 2016) pour une composante de 18-59 ans du Nord, est un minimum de 733 euros par mois (équivalent à 8 800 euros par an net). Dans ce travail, pour ArchiMEDe, les données ont utilisé le seuil de 10 mille euros bruts.

En plus des réseaux éducatifs, professionnels et sociaux, grâce aux archives, il est possible de considérer le réseau familial, en considérant l'équivalent du revenu familial par habitant comme un indicateur de la protection du réseau familial ; différentes structures de revenu peuvent être observées pour les différents groupes considérés, avec des différences possibles par rapport à la situation du revenu individuel seulement (**figure 4**). La famille semble être un réseau efficace de protection contre la vulnérabilité pour les jeunes, en fait, compte tenu du revenu familial équivalent par habitant, toutes les catégories voient une amélioration du positionnement par rapport au seuil de 10 mille, avec des effets plus marqués pour les étudiants (confirmation de l'appartenance à des familles qui peuvent soutenir les jeunes), et significatif pour Q-NEET exTravailleurs et Q-NEET exNEET. En effet, même la catégorie la plus vulnérable, celle du NIEET, améliore sa situation si elle est insérée dans un contexte familial, bien que - comme le montre la figure - elle continue d'être dans une situation de grande faiblesse par rapport aux autres.

Figure 4 - Jeunes de 15 à 29 ans, selon l'appartenance au système formation-travail et le revenu familial équivalent par habitant, Lombardie, 2012 (Box plot)



En 2012 en Lombardie, si l'on considère les jeunes NIEET et Q-NEET comme le groupe de personnes vulnérables, considéré les réseaux de protection des systèmes de formation et de travail, des données d'intérêt émergent.

40 % des personnes vulnérables étaient à l'origine des non-NEET (**tableau 3**). De plus, ce chiffre passe à 44 % si l'on exclut les personnes vulnérables vivant avec leurs parents.

La part des étrangers vulnérables est de 28%, soit presque le double de la part des jeunes étrangers en Lombardie.

En ce qui concerne la citoyenneté, si, d'une part, seuls 28% des étrangers vulnérables vivent avec leurs parents au sein d'une famille (**tableau 4**), d'autre part, la situation inverse est observée pour les Italiens (79% vivent avec leurs parents).

De ces premières observations, il semble qu'il existe des sous-populations avec des vulnérabilités et des réseaux de protection différents. Il est donc nécessaire de mieux identifier les vulnérabilités spécifiques d'une part et, d'autre part, toute politique différente.

Tableau 3 - NIEET et Q-NEET, par type de ménage, Lombardie, 2012 (composition %)

	Autres liens		Total
	familiaux	Fils	
NIEET	45	49	47
Q-NEET exNEET	11	14	13
Q-NEET exTravailleurs	44	37	40
Total NIEET et Q-NEET	100	100	100

Tableau 4 - NIEET et Q-NEET, par citoyenneté et type de famille, Lombardie, 2012 (composition en %)

	Italiens			Étrangers		
	Autres liens		Total	Autres liens		Total
	familiaux	Fils		familiaux	Fils	
NIEET	18	82	100	66	34	100
Q-NEET exNEET	21	79	100	78	22	100
Q-NEET exTravailleurs	24	76	100	79	21	100
Total NIEET et Q-NEET	21	79	100	72	28	100

1.3. Profils de vulnérabilité et politiques

Une fois définies les nouvelles catégories conceptuelles, les données, séparées pour NIEET et Quasi-NEET, peuvent être observées du point de vue de la vulnérabilité économique des familles et de l'importance de certaines variables individuelles d'intérêt (par exemple, le sexe, le type de famille, la citoyenneté,...) dans le but de définir des profils idéaux-typiques au sein des deux sous-groupes.

La finalité de l'analyse est d'observer le jeune en tant qu'élément du système familial. L'extrême richesse de la base de données permet de considérer conjointement la dimension individuelle du jeune et la dimension familiale dans laquelle le jeune est inséré avec l'idée que les politiques d'intervention doivent être modulées selon que les vulnérabilités individuelles s'ajoutent ou non aux vulnérabilités familiales, surtout dans le contexte italien caractérisé par un modèle de *welfare* de type familial.

D'un point de vue purement exploratoire, l'une des méthodes récursives les plus connues utilisées dans la littérature a été utilisée pour le profilage : le *Chi-Squared Automatic Interaction Detection* (CHAID) proposée par Kass (1980). Il s'agit d'une méthode qui, bien qu'elle se réfère aux méthodes plus traditionnelles d'analyse de groupe (p. ex. l'analyse de groupe), diffère de cette dernière en termes d'objectif, de règles de partition et de présence d'une variable *Target* qui représente la variable dépendante du modèle. Essentiellement, la méthode CHAID construit un modèle qui aide à déterminer quelles variables indépendantes considérées expliquent mieux la variable dépendante.

Le modèle est construit à travers des partitions récursives successives d'unités qui sont divisées en sous-groupes progressivement plus homogènes en leur intérieur. Le critère de division est fonction des variables indépendantes considérées, tandis que l'homogénéité au sein des groupes est mesurée sur la variable dépendante. La représentation des segments identifiés se fait à l'aide de l'outil graphique des arbres de décision, dont la structure dépend de la cardinalité des taxonomies des variables à plus grand pouvoir prédictif et de l'agrégation des modalités dont les feuilles définissent les profils (Blangiardo, Mirabelli et Verrecchia, 2010).

En ce qui concerne les règles de partition, qui peuvent être déduites de la structure de l'arbre, la segmentation se fait en considérant la variable prédictive maximale pour chaque branche de l'arbre, contrairement aux méthodes d'analyse des groupes où l'ensemble des variables indépendantes disponibles est utilisé.

La variable dépendante considérée est la vulnérabilité familiale opérationnalisée comme le revenu familial équivalent par habitant. Les ménages dont le revenu brut par habitant est inférieur ou égal à 10 000 € sont considérés comme vulnérables, et non vulnérables les autres.

La variable dépendante considérée est la vulnérabilité familiale opérationnalisée comme le revenu familial équivalent par habitant. Les ménages dont le revenu brut par habitant est inférieur ou égal à 10 000 € sont considérés comme vulnérables, et non les autres.

Afin de définir les profils, certaines caractéristiques des individus et la position du jeune dans la structure familiale ont été observées. Concrètement, nous avons considéré l'âge exprimé en années, le sexe, la citoyenneté (italienne ou non) et le diplôme codifié de trois façons : I. jusqu'au diplôme primaire ; II. diplôme d'études secondaires ; III. diplôme universitaire. La position du jeune dans la structure familiale est une variable qualitative qui est divisée en quatre modes : I. fils ; II. personne en couple; III. famille monoparentale; IV. autre (dans cette catégorie, l'étude de cas la plus courante est celle de la personne qui vit seule).

L'arbre de décision a été limité au premier nœud afin d'obtenir des profils différents pour les deux catégories de NIEET et Quasi-NEET. Les profils qui émergent peuvent être retracés essentiellement jusqu'à deux macro-catégories :

I. Vulnérable dans les situations de vulnérabilité, c'est-à-dire NIEET et Quasi-NEET vivant dans des ménages vulnérables ;

II. vulnérables en l'absence de vulnérabilité familiale, c'est-à-dire NIEET et Quasi-NEET vivant dans des ménages non vulnérables.

Dans la première catégorie, les fragilités individuelles et familiales sont additionnées, ce qui signifie qu'il n'y a pas de réseaux et pas de protection familiale. Dans le second cas, l'absence de réseaux fait partie d'un environnement familial qui protège et éventuellement soutient.

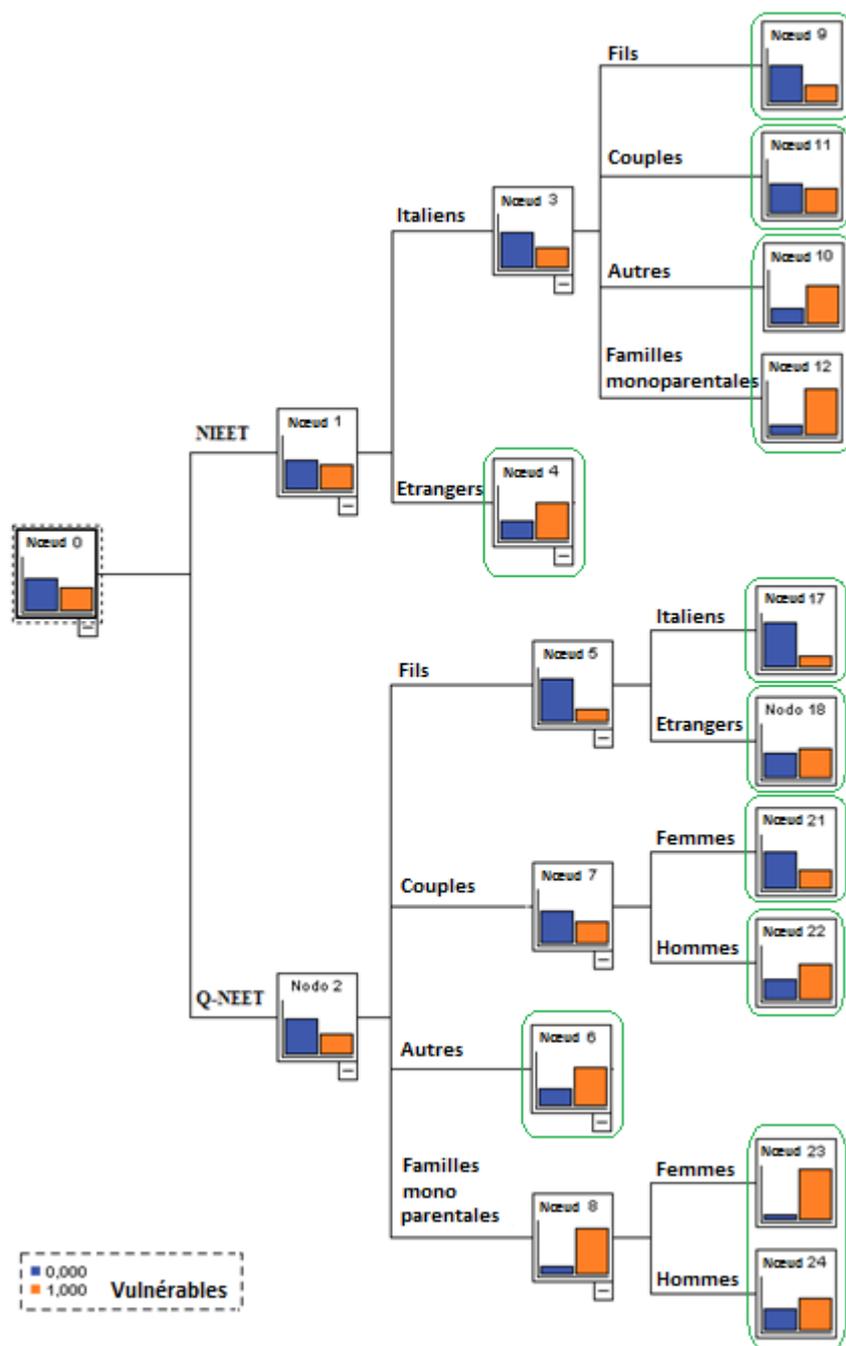
Si l'on regarde les données, tenu compte de l'âge des jeunes et de leurs enfants, non loin des attentes, le profil caractérisé par le plus grand risque de double vulnérabilité est celle de la famille monoparentale. Un risque élevé est également ressenti par la composante étrangère, en particulier parmi les NIEET. L'observation des jeunes en couple, notamment parmi les Quasi-NEET, rappelle une reproduction du modèle familial traditionnel : le risque de double fragilité est élevé, surtout lorsque le jeune homme est dans une situation de vulnérabilité individuelle, mais il est beaucoup plus contenu si c'est la jeune femme qui se trouve dans une telle situation. En général, la condition de fils se caractérise par une plus grande protection contre la fragilité multiple, ce qui est particulièrement vrai pour les jeunes Italiens. Pour les jeunes étrangers, il semble que l'avantage d'être fils ne compense que partiellement le désavantage d'être étranger (**figure 5**).

Si l'on analyse la distribution du revenu brut équivalent par habitant des ménages, on constate que les profils les plus exposés au risque de double vulnérabilité sont aussi ceux qui connaissent les situations de fragilité familiale les plus graves et, inversement, les profils les moins exposés sont ceux qui bénéficient des situations les plus riches (**figure 6**). Le long d'une échelle hypothétique de vulnérabilité familiale des sujets considérés, on trouve, d'une part, les jeunes parent célibataire ou, en tout cas, les jeunes non fils et non en couples, les jeunes hommes Quasi-NEET qui vit en couple et les étrangers, des autres parties, on trouve les fils italiens, qui représentent ensemble plus de la moitié des jeunes NIEET et Quasi-NEET de la population observée (**Figure 7**).

L'observation de la vulnérabilité des jeunes en termes de participation au marché du travail, de participation à des cours de formation et de capacité de revenu d'un point de vue familial plus

général met en évidence le fait que le phénomène n'est pas monolithique. En examinant NEET et Quasi-NEET par rapport à la capacité de revenu familial, des profils se dégagent, caractérisés par une exposition différente au risque de double fragilité et caractérisés par différents degrés de vulnérabilité qui posent des défis plus complexes au décideur public. Les différents besoins exprimés nécessitent des politiques d'intervention différentes qui doivent être modulées en fonction de leur fonction. Pour les jeunes vivant dans leur famille d'origine surtout s'ils sont italiens, des politiques actives spécifiques du marché du travail pourraient suffire. Alors que pour d'autres groupes de jeunes plus vulnérables, il peut être essentiel de combiner les politiques de soutien du revenu avec les politiques d'intégration, en particulier pour la composante étrangère.

Figure 5 - Jeunes de 15 à 29 ans, selon le revenu familial équivalent par habitant, Lombardie 2012 (Arbre de décision)



Note : Orange : Vulnérables ; Bleu : Non vulnérables

Figure 6 - Jeunes de 15-29 ans NIEET et Q-NEET, selon le revenu du ménage par habitant, Lombardie 2012 (Box plot)

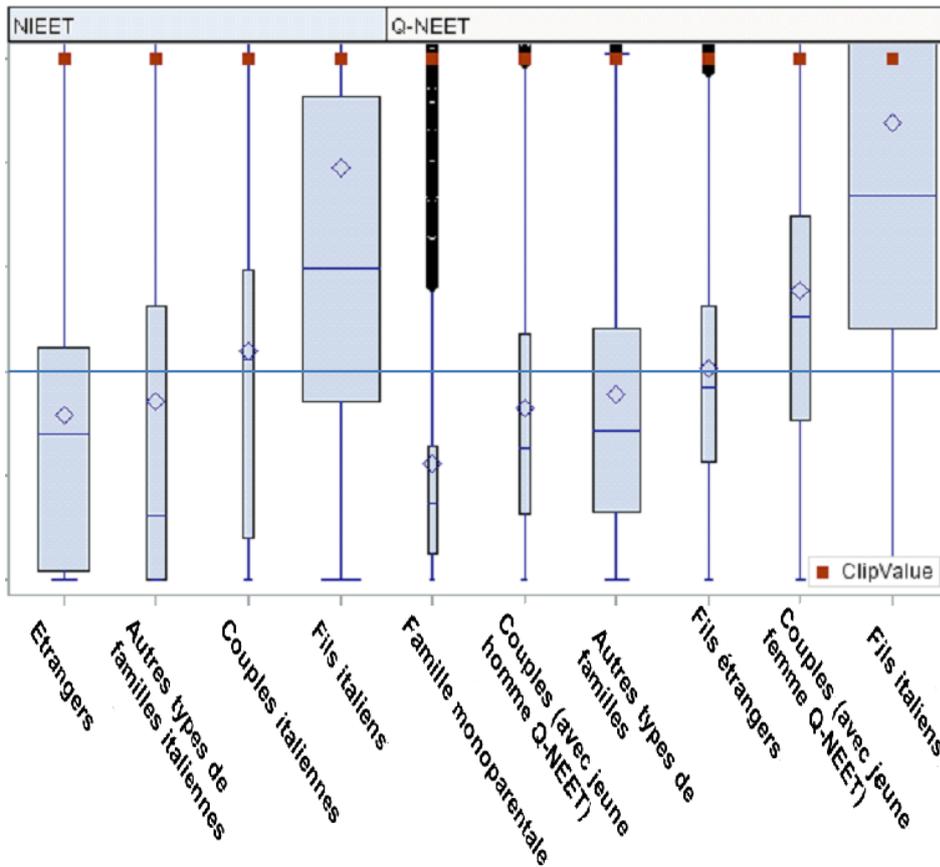


Figure 7 - Jeunes de 15-29 ans NIEET et Q-NEET, selon le revenu médian du ménage par habitant, Lombardie, 2012



5. Conclusions

L'étude du phénomène NEET proposée dans ce document souligne qu'avec les données administratives, il est possible de " reproduire " des indicateurs définis dans d'autres contextes, par exemple dans des enquêtes par sondage.

Le principal avantage des données d'archives, en plus de la réduction du *statistical burden* et des coûts, est de permettre une granularité au niveau sous-régional, d'une utilité particulière pour les décideurs locaux. Mais ce n'est pas tout. En tant que microdonnées, ce type de données permet d'identifier et de concentrer les analyses sur des sous-populations spécifiques d'intérêt.

De plus, la richesse et la multidimensionnalité des données d'archives - surtout si elles résultent de l'intégration de sources administratives multiples, comme dans le cas d'Archimède - offre la possibilité d'expérimenter de nouvelles définitions opérationnelles. En particulier, comme nous l'avons vu, l'inclusion du revenu comme dimension supplémentaire, en plus de celles traditionnellement considérées, permet de segmenter la population des NEET et de délimiter deux autres catégories conceptuelles : NIEET et Quasi-NEET. Il s'agit de catégories de jeunes ayant des profils similaires mais surtout caractérisés par une vulnérabilité sociale et matérielle analogue. En particulier, les Quasi-NEET représentent une catégorie de pertinence car ils sont liés à un phénomène émergent, qui semble être lié à la nouvelle dynamique du marché du travail et aux nouveaux horizons du travail secondaire ou intermittent.

Avec une valeur heuristique et théorique spécifique, la prise en compte du revenu individuel des jeunes nous a permis de faire la lumière non seulement, comme dans la définition classique, sur la présence ou l'absence de réseaux professionnels et de formation, mais aussi sur la présence ou l'absence de réseaux de protection sociale à partir desquels le jeune peut être plus ou moins soutenu. Enfin, l'inclusion du revenu familial équivalent par habitant nous a permis de réfléchir sur le rôle du réseau familial, qui est particulièrement important dans un système de protection sociale comme celui de l'Italie.

Avec l'observation conjointe de la vulnérabilité individuelle des NIEET et des Quasi-NEET et de la famille, il a été possible de délimiter des profils particulièrement faibles en termes de réseaux de soutien. Cette perspective, en particulier, apporte un éclairage nouveau sur le phénomène de la vulnérabilité individuelle des jeunes. L'étude montre qu'en considérant les dimensions traditionnelles du phénomène, il y a un risque d'exclure une " tranche " significative de jeunes vulnérables. Mais il montre aussi que le phénomène de la vulnérabilité des jeunes peut être un défi complexe pour les décideurs politiques - tant en termes de priorités qu'en termes de substance - car il s'agit d'un phénomène qui est loin d'être homogène.

Comme on l'a souligné, si pour les jeunes NIEET et Quasi-NEET vivant dans la famille, des politiques de travail actif ou de formation spécifiques peuvent être suffisantes. Pour d'autres catégories de jeunes exposés à des fragilités multiples et qui ne bénéficient pas du réseau de protection familiale (monoparentaux, étrangers,...), il peut être essentiel de combiner ces politiques avec des mesures de soutien ciblées - au niveau de revenus mais pas seulement - et de politiques d'intégration, en particulier pour les jeunes étrangers.

Bien que conscient des limites actuelles de l'information administrative en Lombardie sur les individus (pour plus de détails voir Ballabio *et al.*, 2018), il est considéré comme intéressant de produire - même en itinérance - des statistiques expérimentales étant donné l'importance que celles-ci ont et pourraient avoir à l'avenir pour les gouvernements des territoires. En ce sens, une fois que les profils de plus grande vulnérabilité parmi la population cible ont été définis, comme cela s'est produit dans ce travail, la richesse et le détail des données permet également l'analyse de la distribution des profils spécifiques d'intérêt sur le territoire pour une allocation plus efficace des ressources également au niveau sous-régional.

Bibliographie

- [1] Alwang J., Siegel P.B., Jorgensen S.L., «Vulnerability. A View from Different Disciplines», Social Protection Discussion Paper, n° 0015, The World Bank, Washington, D.C., 2001.
- [2] Ambroselli, S., «I codici identificativi univoci all'interno del SIM (Sistema Integrato di Microdati)», Istat Working Paper, n° 5, 2015.
- [3] Ballabio S., Carra A., Casacci S., Ferrazza D., Verrecchia F., Vitalini A., Viviano L.C., «Quasi NEET? Un'analisi a partire dai dati ArchIMEDe», Neeting2. Convegno internazionale sui NEET, Milano, dicembre 1, 2017.
- [4] Ballabio S., Carra A., Casacci S., Ferrazza D., Verrecchia F., Vitalini A., Viviano L.C., «Local decisions and new guidelines of the Official Statistics. An experiment with Early School Leavers to measure the quality of integrated administrative data used to support local policies», Q2018, Kraków, June 26-29, 2018.
- [5] Bebbington, A., «Capitals and Capabilities: A Framework for Analyzing Peasant Viability, Rural Livelihoods and Poverty», World Development, vol. 27, pp. 2021-2034, 1999.
- [6] Blangiardo, G.C., Mirabelli, S.M. e Verrecchia, F., «L'immigrato elettore: idealtipi a confronto», Rivista Italiana di Economia Demografia e Statistica, vol. LXIV, n. 4, pp. 63-70, 2010.
- [7] Caritas, «Immigrazione. Dossier Statistico 2002: XII Rapporto sull'immigrazione Casitas – Migrantes», Nuova Anterem, Pomezia, 2002.
- [8] Esping-Andersen, G. «The three worlds of welfare capitalism», Princeton, New Jersey, Princeton University Press, 1990.
- [9] European Commission, «La garanzia per i giovani e l'iniziativa a favore dell'occupazione giovanile a tre anni di distanza», Comunicazione, 2016. On-line: <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1477901398883&uri=CELEX:52016DC0646>, download mai 2018.
- [10] Ferrera M., «Le politiche sociali. L'Italia in una prospettiva comparata», Mulino, 2006.
- [11] Freyssinet J., «How can social vulnerability be measured : a work in progress», The 3rd OECD World Forum on Statistics, Knowledge and Policy, Busan, October 27-30, 2009.
- [12] Garofalo, G., «Il Progetto ARCHIMEDE obiettivi e risultati sperimentali», Istat Working Paper, n° 9, 2014.
- [13] Istat, «Censimento in pillole - Lombardia», comunicato stampa, 30 gennaio 2013. On-line: <https://www.istat.it/it/files//2013/01/Censimento-in-pillole-Lombardia.pdf>, download mai 2018.
- [14] Istat, «L'Italia del censimento. Struttura demografica e processo di rilevazione. Lombardia», Istat, Roma, 2013. On-line: https://www.istat.it/it/files/2013/01/Lombardia_completo.pdf, download mai 2018.
- [15] Istat, «9° Censimento dell'industria e dei servizi e Censimento delle istituzioni non profit. Principali risultati e processo di rilevazione. Lombardia», Istat, Roma, 2014.
- [16] Istat, «14° Censimento Generale della Popolazione e delle Abitazioni. Censimento 2001». On-line: <http://dawinci.istat.it/MD/index.html>, download mai 2018.
- [17] Istat, «Censimento Popolazione Abitazioni». On-line: <http://pop-dati2.istat.it/Index.aspx>, download mai 2018.
- [18] Kass, G.V. (1980), «An exploratory technique for investigating large quantities of categorical data», Applied statistics, vol. 29, pp. 119-127.

[19] Morrone A., Scrivens K., Smith C., Balestra C., « Measuring vulnerability and resilience in OECD countries », IARIW-OECD Conference on Economic Insecurity, Paris, November 22–23, 2011.

[20] Moser C., «The Asset Vulnerability Framework: Reassessing Urban Poverty Reduction Strategies», World Development, vol. 26, n° 1, pp. 1-19, 1998.

[21] Moser C., Holland J., «Household Responses to Poverty and Vulnerability - Volume 4: Confronting Crisis in Cawama, Lusaka, Zambia», Urban Management Programme, n° 24, The World Bank, Washington, D.C., 1997.

[22] Reardon T., Vosti S., «Links Between Rural Poverty and the Environment in Developing Countries: Asset Categories and Investment Poverty », World Development, vol. 23 n° 9, pp. 1495-1506, 1995.

[23] Statistics Canada, «Survey Methods and Practices», Catalogue no. 12-587-X, 2010. On-line <http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf>., download mai 2018.